Mouse labial-like homeobox-containing genes : Structure and expression during embryogenesis

Paula Murphy

PhD

University of Edinburgh

1991

Declaration

I declare that this thesis was composed by myself. Contributions of others to the work are clearly indicated.

Quote

"Mouse development is like turning a lump of iron ore into a space shuttle overnight"

Miroslav Holub

poet and immunologist

BBC radio 4, 1990.

Acknowledgements.

This project was instigated and supervised by Dr. Robert Hill to whom I am grateful, not only for his constant help and support but also for making the experience a pleasant one. I also especially thank Dr. Duncan Davidson for sharing his time, his interest in developmental biology and for familiarising me with the mouse embryo. I thank Dr. Nick Hastie for the large part he played in making it possible for me to study with the MRC.

The work presented in chapter 5 was carried out in collaboration with Dr. Gillian M. Morriss-Kay in the Human Anatomy Department at Oxford University. The embryonic material for this work was prepared by Gillian. I also thank her for useful discussions and for critically reading the relevant part of the text.

Some of the work presented in this thesis has been published in Murphy *et al.*, 1989 and Murphy & Hill, 1991. Reprints are included in the appendix.

There are many people in the MRC unit who deserve my gratitude for making the work easier or more pleasant. Maureen and I started and finished our projects together and we shared the traumas and joys of a PhD. My fellow country-women and friends, Paula and Sinead, helped me keep up the side, especially during the world cup. I thank all the other members of the lab, too many to mention individually, for being such good company. I especially thank Chris for all her help and for running the place gently but effectively. Thanks also to Chris and Sheila for providing all the goodies and for creating the illusion that enzymes grow in freezers. Thanks to Liz for her organisation of the *in situ* lab and, together with Duncan, for teaching me histology. I thank Bill for the whisky tasting, where I discovered Lagavulin, and for organising the skittles outings.

The photographic and library services at the unit were particularly valuable in the production of this thesis. I thank Norman, Sandy and Douglas in the photography department and the librarians, Sheila and Helen.

Edinburgh is a beautiful place and I have enjoyed discovering something of Scottish life, but most important were the people I met and the friendships I made, both inside and outside the lab. Among these were the best bunch of women in the world and special friends; Moire, Delia, Sophie, Pauline and Cathy. I thank Tarana and Steve for the company and the food. I thank Maureen and Dave for some good nights and memorable conversations.

John kept me off the streets in the penniless days between PhD. and postdoc. I thank him for that, for proof-reading the thesis and for understanding the frustrations of thesis writing.

My family, as always, were very important during this endeavour and visits to and from Ireland were a very welcome distraction. My siblings; Dermot, Maria, Betty, Bernie, Seamus, Sean and Michael, have each been very supportive in their own way. I especially thank my father, James, who always gave me everything he could. This thesis is dedicated to my father and to my mother Mary, who sowed the seeds of achievement.

Abstract

While mouse development has been well described at a morphological level, very little is known about how development is regulated. In contrast, the ease of developmental analysis in *Drosophila* has led to the identification of a large number of developmentally important genes. Molecular characterisation revealed that many of the genes involved specifically in determining the *Drosophila* body plan contain a conserved sequence called the homeobox. This sequence is highly conserved through evolution and so it can be used to isolate homologous genes in other species. In this way more than 40 homeobox-containing genes have been identified in the mouse. The high level of sequence conservation and the temporally and spatially restricted expression of the mouse genes during development indicate that they are also developmental regulators involved in conferring spatial information within the embryo. Thus, through knowledge of *Drosophila* development and the techniques of molecular biology, it is now possible to study mouse developmental genes in detail.

In this thesis, the characterisation of two mouse homeobox-containing genes, *Hox 2.9* and *Hox 1.6*, is presented. Sequence analysis revealed that these genes are closely related and that, among *Drosophila* genes, they are most similar to *labial* in the *Antennapedia* complex. They are therefore thought to have arisen by duplication of a single ancestral gene. As well as being structurally similar the genes share many features of their expression patterns. Both genes are expressed early in development ($7^1/_2$ days) and, unlike other known mouse homeobox-containing genes, they are not expressed after 11 days of development. At 8 days the genes share the same anterior boundary of expression in the hindbrain and in the later embryo, with the exception of persistent *Hox 2.9* expression in the hindbrain, they have the same anterioposterior restrictions. This indicates that the genes are functionally similar and also that they respond to at least some of the same signals in the embryo.

A striking difference between the expression patterns of mouse *labial*-like genes is the unique expression of *Hox 2.9* in a single segmental unit (rhombomere 4) of the hindbrain (from $8^1/_2$ days). This expression coincides perfectly with the morphological extent of rhombomere 4 and persists throughout the period that rhombomeres are visible (up to 11 days). It is therefore suggested that *Hox 2.9* participates in conferring segment identity. In addition neural crest cells that arise from rhombomere 4 specifically express *Hox 2.9* and this supports the idea of neural crest cells being patterned according to their position of origin in the central nervous system. Detailed analysis of the onset of segmental expression of *Hox 2.9* and another segmentally expressed gene in the hindbrain, *Krox 20*, showed that *Hox 2.9* expression becomes localised from a broad domain at $8^1/_2$ days of development, up to 6 hours before rhombomeres are clearly visible.

Retinoic acid is a strong candidate for a natural morphogen in the vertebrate embryo. The effect of *in vitro* treatment with retinoic acid on segmentation of the mouse hindbrain and on the expression of *Hox 2.9* and *Krox 20* was therefore analysed. It was found that segmentation in treated embryos is abnormal and that the clear segmental localisation of expression of the two genes is not found. The hindbrain expression domains are shifted rostrally following treatment and while the expression of the two genes remains mutually exclusive there is no longer a single planar boundary between the domains. Instead there is an irregular alternation of cells expressing the two genes at the boundary.

Two differential splicing products of *Hox 1.6* were isolated from the developing embryo. A comparison was made of the distribution of these transcripts, only one of which can code for a homeodomain containing protein. It was found that the relative proportion of homeodomain producing message decreases as development proceeds.

Table of Contents

Chapter 1

Introduction

This thesis is concerned with the activity of homeobox-containing genes during early mouse development. Before pursuing in detail the issues of the thesis, a number of areas need to be reviewed. The introduction is therefore divided into four sections. Section 1 describes the processes involved in early mouse development. Section 2 reviews current knowledge on the regulation of early development in the mouse and other vertebrates. Section 3 is concerned with the role of homeobox-containing genes in *Drosophila*, where these genes were first isolated and are best understood. Finally, section 4 deals with homeobox-containing genes in the mouse.

1.1 Early embryonic development.

It takes 20 days for the mouse embryo to develop *in utero*. The normal process produces a perfectly formed individual from a single fertilised egg cell, with hundreds of cell and tissue types operating in harmony to carry out all the necessary functions of this complex organism. During the first 10 days the basic body plan is established with major structural subdivisions and initial differentiation of the major body organs. During later stages, specific organs and tissues are formed in preparation for independent life. This thesis is largely concerned with the early stages of development and the acquisition of positional identity by the cells within the embryo. The processes involved in early development are therefore described in this section (general references; Hogan *et al.*, 1986; Theiler, 1972; Rugh, 1990; Jackson, 1989).

In the mouse, the initial divisions of the fertilised egg are slow and the cells form a loose aggregate in which individual cells are not developmentally restricted (up to the 8-cell stage). This was demonstrated by the fact that chimeric mice produced from two 8-cell embryos have cells from both donors in all embryonic and extraembryonic tissues (McLaren, 1976; Kelly, 1977). Compaction of the cells occurs at the 8-cell stage, forming a morula. This is an important event as it allows cell communication through gap junctions for the first time. The cells now obtain

polarity, with an inside face and an outside face, and this is thought to be the basis of the formation of the first two distinctive cell lineages in the later blastocyst (64 cells) (Johnson & Ziomek, 1981).

The blastocyst is hollow and asymmetrical. The inner cells at the thicker side form the inner cell mass (ICM) and those on the outside form the trophectoderm. These two cell types undergo a second differentiation event before the embryo implants. The trophectoderm differentiates according to contact with the ICM; those cells in contact with the ICM form polar trophectoderm while those in contact with the blastocoel form polyploid giant cells. The ICM cells that are in contact with the blastocoel form primitive endoderm while the core cells form primitive ectoderm, or the epiblast. The entire embryo is formed from cells of the epiblast (Gardner, 1982) whereas all other cells form extraembryonic structures such as the yolk sac, placenta and protective membranes.

The blastocyst hatches from the zona pellucida and implants into the uterine wall at $4^1/_2$ days. Until this time there are no nutrients available to the embryo and therefore the embryonic mass does not increase even though the cells divide and undergo important changes. Implantation begins when the trophoblastic giant cells penetrate the epithelial layer of the uterine wall which then proliferates around the blastocyst. Meanwhile the epiblast is pushed down into the blastocoel to form the egg cylinder as the polar trophectoderm proliferates. The primitive endoderm proliferates to line the entire blastocoel where the cells differentiate into parietal and visceral endoderm. A cavity forms within the epiblast turning the egg cylinder into a two layered cup of ectoderm surrounded by endoderm. This is the situation found at 6 days, prior to gastrulation. Gastrulation begins between $6^1/_2$ and 7 days and marks the beginning of the proliferation and differentiation of the embryo proper. Earlier events are largely concerned with the development of supporting structures required for intra-uterine life.

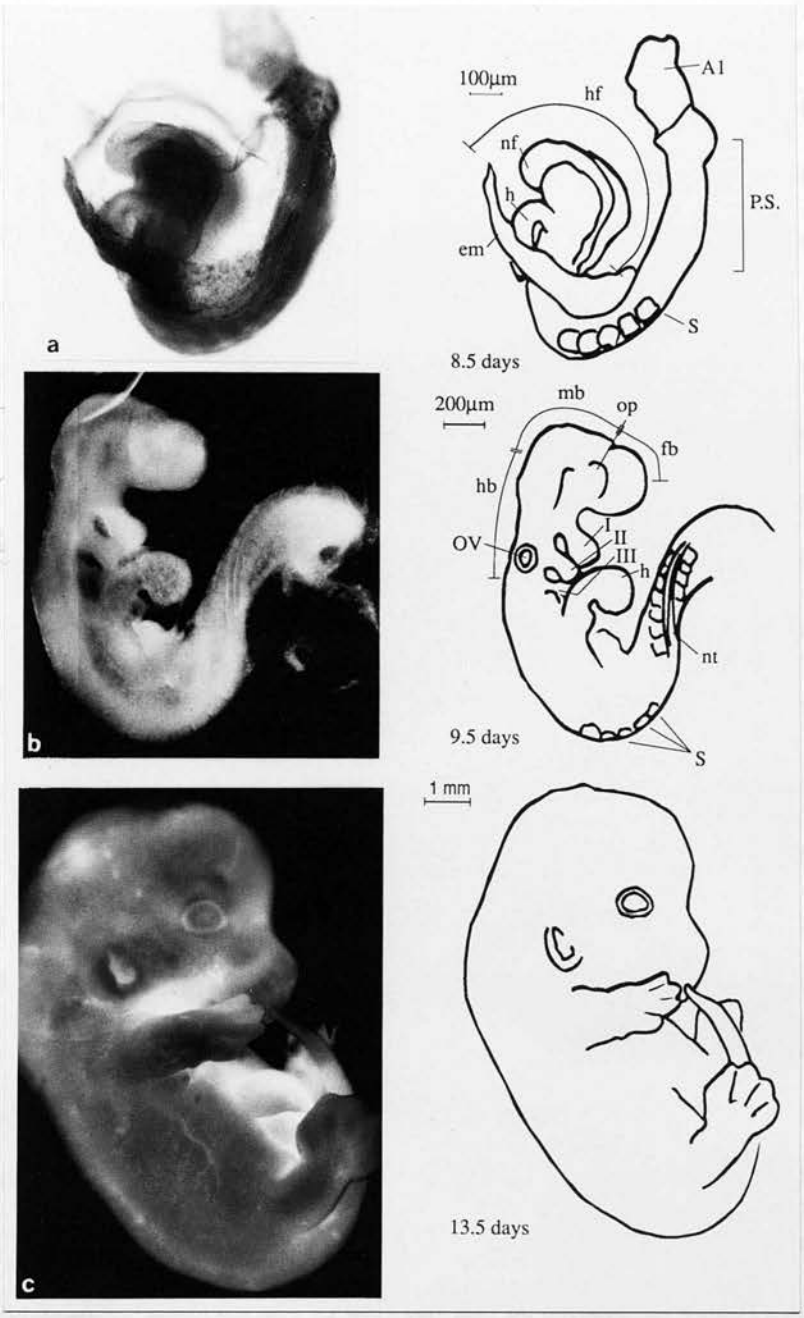


Figure 1.1. The developing mouse embryo. (a) shows a photograph and a schematic diagram of an $8^1/_2$ day embryo with open neural folds in the head fold and at least 6 pairs of somites. (b) shows an embryo that is approximately 24 hours older. Note that the embryo has almost fully rotated 180^o about its own axis, so that it is now ventrally concave (the posterior regions are still in the process of turning). The chambers of the developing brain are now distinguishable. (c) shows an embryo at $13^1/_2$ days with the external features of a recognisable fetus. hf, headfold; nf, neuralfold; h, heart; em, extraembryonic membrane; s, somite; ps, primitive streak; Al, allantois; fb, forebrain; mb, midbrain; hb, hindbrain; op, optic placode; ov, otic vesicle; I, II, III, pharyngeal arches I, II and III; nt, neural tube.

1.1.1 Gastrulation

During gastrulation the three germ layers of the embryo are formed with much cell proliferation and cell movement. The primitive streak appears at the then posterior end of the epiblast (the embryo later elongates posteriorly) giving the embryo polarity for the first time. Cells move through the primitive streak and spread out beneath the ectoderm forming mesoderm. Beneath the anterior tip of the primitive streak (Hensons' node in the chicken) a special set of mesoderm cells differentiate. As gastrulation proceeds the primitive streak moves posteriorly leaving behind a trail of these cells which form the notochord. At about $7^1/_2$ days the headfold begins to form as the ectoderm anterior to the primitive streak thickens and, together with the underlying mesoderm, which has moved anteriorly from the primitive streak, is raised up by the fore-gut pocket which indents ventrally. The fore-gut pocket is lined with endoderm which is also thought to have originated from the primitive ectoderm during gastrulation.

The embryo acquires an anterioposterior (AP) axis and a dorsoventral (DV) axis during gastrulation. The primitive streak is thought to have the same organisational role as the dorsal blastopore lip has in amphibians (Hamburger, 1988) during this process. However, this has not been clearly demonstrated, due largely to the relative inaccessibility of the mouse embryo at these stages. At $7^1/_2$ days the mouse embryo measures only about 250μm x 200μm and is embedded in a decidual swelling approximately 7 times its size. It is technically difficult to dissect the embryo from the decidual swelling without damage and culturing 7-8 day embryos is therefore particularly difficult (Tam & Snow, 1980). Preimplantation blastocysts can be cultured up to, but not beyond, the time of implantation. Embryos at $8^1/_2$ days can also be successfully cultured for several days because they are larger and sturdier. Manipulating the early embryo therefore presents a huge problem to the embryologist and the determinitive events that occur during gastrulation remain largely a mystery. Despite the difficulties, many studies have been carried out on cell determination (Snow, 1981; Beddington, 1982; reviewed in Beddington, 1987) and at least have demonstrated the central importance of gastrulation. No heterogeneity has been detected within the epiblast prior to gastrulation (Beddington, 1983) but during and following gastrulation the

subdivision of the embryo becomes clear. This is reflected in the structural variations along the AP and DV axes and also in the first localised expression of specific genes, most notably homeobox-containing genes, along the AP axis (Gaunt, 1987; Holland & Hogan, 1988b).

Gastrulation continues in posterior regions until about $9^1/_2$ days as the embryo continues to elongate. Over this period the relatively simple structure of the early embryo is transformed into a much more complex entity displaying distinctive features of the fetus (figure 1.1).

1.1.2 Segmentation of the mesoderm.

After invagination, the mesoderm that lies adjacent and parallel to the notochord (paraxial mesoderm) condenses into paired blocks of cells called somites. This is a progressive process beginning just behind the headfold between $7^3/_4$ and 8 days and proceeding posteriorly. Between the most recently formed pair of somites and the primitive streak lies paraxial mesoderm that has not yet condensed (presomitic mesoderm). Close examination of this region using scanning electron miscroscopy revealed that a prepattern of somitomeres exists before overt segmentation in the earliest primitive streak stage embryo (Tam *et al.*, 1982).

Approximately 65 pairs of somites form in the mouse. Initially they are said to have an epithelial structure since they are surrounded by a layer of extracellular matrix glycoproteins (Leivo *et al.*, 1980). Later the somites are subdivided into three regions: (1) The sclerotome, which is involved in the formation of vertebrae around the neural tube and notochord (anterior sclerotome cells of one pair of somites join with posterior sclerotome cells of the adjacent pair to form a single vertebra). (2) The dermatome, which forms connective tissue and dermis of the skin. (3) The myotome, which differentiates into muscle.

The vertebrate mesodermal layer is not entirely segmented. In the head no somites form but there is evidence for seven somitomeres (Meier & Tam, 1982). Outside the paraxial mesoderm, the developing kidney is segmented but the rest of the lateral lying mesoderm (lateral

plate mesoderm) is not. There are however obvious regional differences within the lateral plate and since the mesoderm of the more primitive vertebrate, *Amphioxus*, is completely segmented, incomplete segmentation is likely to be an evolutionary modification (Goodrich, 1985).

Manipulation of somites is very difficult in the mouse (for the reasons mentioned previously) but experiments in the chicken have shown that the particular fate of individual somites is determined by the position and/or the time at which they are formed during the gastrulation process (reviewed by Hogan *et al.*, 1985). This was shown by transplantation experiments in which transplanted somites formed structures appropriate to their origin (Kieny *et al.*, 1972).

1.1.3 Neurulation

One of the most important aspects of gastrulation is thought to be the bringing together of different tissue types in the correct arrangement for subsequent inductive interactions. There is some evidence that during gastrulation the embryo remains developmentally labile and that final differentiation depends on position within the embryo (Beddington, 1982) and on tissue interactions (Snow, 1981). One of the processes in which such tissue interactions are vital is neurulation. The notochord, running along the mid-line of the AP axis of the embryo, induces overlying ectoderm to form neuroectoderm. This tissue is the precursor of the central nervous system (CNS) and the neural crest cells.

The neuroectoderm first appears as a flat monolayer of cells anterior to the primitive streak as it retreats. This early structure is called the neural plate. Some of the major structural subdivisions of the CNS can already be seen in the neural plate. However, it is not known if this reflects commitment of the cells to the formation of specific structures or if it is simply a consequence of the shape of the underlying notochord (Morriss-Kay, 1981). A more reliable indication of regionalisation is the expression of specific genes.

The lateral edges of the neural plate fold upward to form the neural groove. The folds become more exaggerated until the edges fuse mid-dorsally. In this process neuroectoderm fuses with neuroectoderm and non neuroectoderm fuses with non neuroectoderm to enclose the neural tube. This process begins just posterior to the head fold and progresses anteriorly and posteriorly. Neural fold closure is immediately preceeded by emigration of the neural crest cells from the lateral edges of the folds. Some of the neural crest cells contribute to the sensory ganglia that form outside the CNS. Others are non-neuronal and play important roles in contributing to the formation of many structures. These will be more fully described in later sections (section 1.2.4, chapter 5).

The brain forms from the anterior part of the neural tube which is initially divided into three major chambers; the forebrain, midbrain and hindbrain. The forebrain subsequently forms the cerebral cortex, the basal ganglia and the thalamus. The midbrain, together with the anterior hindbrain (metencephalon), forms the cerebellum and the pons. The posterior hindbrain (myelencephalon) forms the medulla oblongata. The posterior neural tube is more uniform in shape and forms the spinal chord. Neuroblasts, the primitive nerve cells, are first seen just after the neural folds close (approximately 9 days). Different classes of neural cells subsequently differentiate at fixed locations along the DV axis. Motor neurons form at the ventral side (basal plates) and sensory nerves form at the dorsal side (alar plates). The process of neuronal differentiation continues throughout the later stages of development as a complex network of axons and synaptic contacts are established to facilitate communication throughout the body.

Little is known about how the network of nerves is established. At the ventral midline of the neural tube there is a distinctive region called the floor plate which appears to play a role in the patterning of axonal growth from at least one type of sensory neuron, the commissural neuron. Axons extend from commissural neurons toward the ventral side of the neural tube, through the motor column and into the floor plate. Co-culturing experiments have indicated that the cells of the floor plate release a chemotactic factor that guides commissural axons (Tessier-Lavigne *et al.*, 1988). Floor plate cells were also shown to have a polarising effect (section 1.2.6) when

transplanted to the developing chicken limb bud (Wagner *et al.*, 1990) indicating that they are the source of a substance capable of disturbing the pattern of limb formation in the same way as retinoic acid. The distinctive character of the floor plate was also shown by the fact that it forms a separate compartment in the developing chicken hindbrain (Fraser *et al.*, 1990).

1.1.4 Segmentation of the central nervous system.

The major AP subdivisions of the CNS, into the chambers of the brain and the trunk neural tube have already been described. Further subdivisions become obvious after $8^1/_2$ days with the appearance of repetitive undulations along the neural epithelium, called neuromeres. These are most prominent within the hindbrain where they are specifically called rhombomeres (figure 1.2). Neuromeres were first observed in 1828 (Baer v.) but were not described in detail, or named, until the work of Orr in 1887. They were subsequently described in a wide variety of vertebrates (Streeter, 1908; Neal, 1918; Adelman, 1925; Kallen, 1953; Vaage, 1969; Tuckett *et al.*, 1985; Sakai, 1987) but their significance as segmental units was controversial. There were two alternative possibilities, as proposed by Neal (1918): that neuromeres are simply a manifestation of compression and strain within the growing neural tube or that they represent segmental organisation within the developing CNS. Neal and other observers agreed that separate cases could be made for hindbrain rhombomeres and for neuromeres elsewhere.

The prominent nature of the rhombomeres has made them more accessible to detailed examination. Rhombomere formation was studied in the rat (Tuckett *et al.*, 1985) and in the mouse (Sakai, 1987) where it follows a definite pattern with individual structures appearing progressively but not in a linear order. Adelmann (1924) originally suggested that rapid growth in a confined space explained the appearance of rhombomeres, however, this explaination seems too simple. Kallen (1953) detected proliferation centres within the neuroepithelium and showed that they coincided topographically with neuromerical bulges. Tuckett *et al.* (1985) could not correlate mitotic index peaks with visible rhombomeres but suggested that proliferation preceeds

rhombomere appearance. Tuckett and Morriss-Kay (1985) subsequently proposed a model for rhombomere formation based on the detection of microtubules at rhombomere boundaries and microfilaments at the ventricular concave rhombomere surface. They suggested that the neural epithelium bulges between the microtubule blocks, along the line of least resistance, away from the microfilament rich surface.

Rhombomere formation preceeds the outgrowth of motor nerves and the ingrowth of sensory nerve fibres from the ganglia. However, several authors have noted the subsequent correspondence between the position of nerves and rhombomeres (Orr, 1887; Streeter, 1908; Neal, 1918; Adelman, 1924; Vaage, 1969) and this feature was taken by some as evidence to suggest a developmental significance for rhombomeres. However, it was not until Lumsden and Keynes (1989) observed the pattern of nerve formation within the chicken hindbrain (by labelling the neurofilaments) that the segmental basis to neurogenesis and the relevance of rhombomeres was revealed. Neurons first develop in alternate rhombomeres and only subsequently appear in the intervening rhombomeres. Later there is also co-operation between motor axons in adjacent rhombomeres in forming a single motor nerve root, so that motor nerves arise from alternate rhombomeres. Fraser *et al.* (1990) later showed that the rhombomeres represent lineage compartments between which there is no cell mixing once the boundaries have formed. Rhombomeres are therefore units within the embryo through which development is organised. This is further supported by the fact that *Krox 20*, a gene producing a DNA-binding protein and therefore a potential regulatory gene, is expressed segmentally in rhombomeres 3 and 5 of the mouse hindbrain (Wilkinson *et al.*, 1989a).

Although it is now clear that rhombomeres are segments within the developing CNS, the case for trunk neuromeres is not so strong. Their formation proceeds in the same direction as somitogenesis (Sakai, 1987) and there is a one to one correspondence between neuromeres and somites. For this reason it was suggested that trunk neuromeres are simply formed as a result of pressure on the neural tube from the somites (Neal, 1918). However in a lower vertebrate, the zebrafish, the pattern of neurogenesis is similar in the hindbrain and spinal chord with the first

neurons appearing periodically (Hanneman *et al.*, 1988). This suggests that there is a similar segmental pattern throughout the CNS. The original segmental pattern becomes obscured as the complexity of nerves increases and this may be why a similar pattern is not discernable in higher vertebrates. Kimmell *et al.* (1988) report that heat-shock treatment of developing zebrafish embryos disturbs somites and spinal segments co-ordinately. This can be interpreted in two ways. Either segments are formed co-ordinately in both tissues or one of the tissues is not segmented but affected indirectly via the other. The second interpretation is supported by work in the chicken where somites were rotated by 180^{o} with respect to the neural tube (Keynes & Stern, 1984). The motor neurons which subsequently formed, were found to grow out in altered positions in order to enter the correct part of the somite, suggesting that the segmental pattern of spinal nerves is imposed by the somites and does not reflect an intrinsic periodicity.

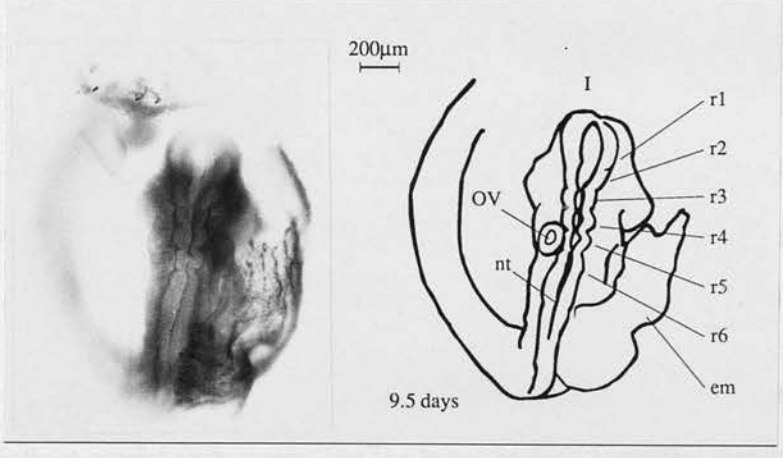


Figure 1.2. A dorsal view of a $9^{1}/_{2}$ day embryo showing well formed rhombomeres in the hindbrain. I, istmus; r1-r6, rhombomeres 1-6; ov, otic vesicle; nt, neural tube; em, extraembryonic membranes.

1.2 The regulation of early development: laying down the body plan.

The differentiation of initially similar cells into diverse and specialised fates is the means by which complex multicellular organisms develop. As early as 1934, T.H. Morgan proposed the 'theory of differential gene activity in development' and we now know that differential expression of sets of genes, under the control of specific regulators, is the basis of differentiation. But the concept of differential gene expression is not sufficient to explain how a three dimensional body plan is organised. Cells must be able to communicate; to transmit and receive information about their relative positions within the developing embryo in order to establish their fate.

There are a number of ways by which such information can be transmitted and here, the various possibilities are explored in the light of current evidence. It appears that a number of mechanisms are in operation. Cells have 'memories' enabling them to receive layers of information as the embryo develops, channeling them toward their final fates. The overall effect is one in which 'a consistent and accurate result is obtained by the combination of several processes' (Gurdon, 1989). It appears that early information is crude and further interactions lead to finer detail, as was shown for the establishment of dorso-ventral position in amphibian mesoderm (Dale & Slack, 1987). We are only now beginning to understand the nature of the information that is passed between cells in an embryo, and how the receiving cells interpret this information. Classical experiments, involving embryo manipulation, have given us a basic understanding of the overall processes and provided good systems in which to test various models.

The large amphibian embryo is more accessible for manipulation than the mammalian embryo, especially at crucial early stages, and has been the most important vertebrate system for studying cell-cell interactions and positional information. Much of the evidence presented here has therefore been established in the amphibian. It is likely that the processes are similar in mammals, although the timing and pattern of events may be different.

In the amphibian, heterogeneity in the egg cytoplasm is utilised to lay down the first rudiments of the body plan: as the egg cleaves the information is unequally apportioned to the daughter cells. However there is no evidence for localised embryonic information in the mammalian egg (Woodland, 1989). The means by which initial positional information is acquired is therefore unknown.

1.2.1 Embryonic Induction.

Induction is the process whereby a subset of similar cells are stimulated to differentiate in response to cells of another kind lying in close proximity. This describes most of the processes of the early embryo and is probably the most important means by which information is passed. Inductive interactions have been described extensively in vertebrates and are also involved in the development of invertebrates such as the nematode *C. elegans* (Sternberg & Horvitz, 1986). It now appears that induction is a general description covering a number of different cell-cell interactions (Gurdon, 1987, for review). The production of a single final cell type, for example liver, may involve several separate inductions during development (Wessells, 1977).

In mammals, induction has been studied during organogenesis, for example during development of the kidney (Saxen, 1987). However, for the reasons mentioned previously, our knowledge of early induction comes mainly from the amphibian system. Three particular inductive responses have been most actively investigated: (1) Mesoderm induction, where animal pole cells of the blastula are induced to form mesoderm by vegetal pole cells (reviewed by Smith, 1989). (2) Neural induction, where mesoderm cells at the dorsal blastopore lip and inside the embryo induce the overlying ectoderm to form nerve cells. (3) Lens induction, where optic lobes of the forebrain induce overlying ectoderm to invaginate and form a lens vessicle (Jacobson, 1966).

Results from these systems and from organ induction in mammals and other vertebrates, have revealed the heterogeneity of the inductive response. Nevertheless few general characteristics

can be described. Firstly, the timing of the inductive event is limited by both the inducing and responding tissues. The capacity to induce or to be induced is only present at the appropriate time (Gurdon, 1989), emphasising the part that induction plays in a series of chronological events. Secondly, it appears that the localisation of response is controlled by the cells that produce the signal since the number of cells that are capable of responding is greater than the number that actually do so. The proximity of the inducing and responding cells is important (Nieuwkoop *et al*, 1952; Gurdon, 1989). However recent evidence suggests that the responding cells produce a secondary signal (the community effect), the concentration of which determines the size and perhaps the nature of the response (Gurdon, 1988). Thirdly, there is a sharp cut-off line between cells which respond and those that do not. This is most simply explained by a threshold effect where an individual cell responds fully or not at all (Slack, 1983). This type of response would be enhanced by 'the community effect' mentioned above. Finally, at least some inductive interactions do not require cell-cell contact (Karkinen-Jaaskelainen, 1978; Grunz & Tacke, 1986) and therefore are transmitted by diffusible factors.

The system of mesoderm induction in the amphibian has been most fruitful in revealing the nature of early cell-cell interactions and recently, candidates for the transmitted signal(s) have been isolated. Initially it was found that before the 64 cell stage in the blastula, isolated animal and vegetal pole cells will only form ectoderm and endoderm, but at later stages mesoderm cells are also formed (Nakamura & Matsuzawa, 1967). It was shown by Nieuwkoop (1969) that mesoderm cells are only produced if animal and vegetal cells are cultured together and also that mesoderm is formed entirely from ectoderm following induction by prospective endoderm (Sundarwati & Nieuwkoop, 1971). Boternbrood & Nieuwkoop (1973) later showed that the vegetal pole cells also determine the type of mesoderm that is formed, with dorsal vegetal cells inducing dorsal mesoderm (notochord and muscle) and ventral vegetal cells inducing mesenchyme and mesothelium. These results however contradict the fate map since ventral vegetal cells induce little or no muscle whereas fate mapping experiments suggest that most of the muscle in the embryo is actually formed from ventral cells. This was theoretically resolved by the proposal of the three

signal model (Smith & Slack, 1983) which involves a signal from dorsal vegetal cells inducing notochord and muscle, a signal from ventral vegetal cells inducing blood and mesenchyme and a third dorsalising signal from the newly induced dorsal mesoderm. The dorsalising signal was proposed to induce adjacent ventral mesoderm to a more dorsal fate. The existence of such a dorsalising factor has been supported by transplantation experiments (Slack & Forman, 1980) and it may be one of the factors involved in Spemanns' organiser graft (Hamburger, 1988) where transplantation of a second dorsal blastopore lip to the ventral marginal zone led to a mirror image, double-dorsal embryo.

Following a long search for mesoderm inducing factors (MIFs) some likely candidates have recently been purified. These belong to the family of protein growth factors and fall into two classes. The first class comprises XTC-MIF (Smith, 1987) and TGFß2 (Rosa *et al.*, 1988) which are related to transforming growth factor type ß (TGFß). The second class contains acidic and basic fibroblast growth factors (FGF) (Slack *et al.*, 1987; Kimelman & Kirshner, 1987; Slack *et al.*, 1988). One important difference between the two classes is that the TGFß class can induce all mesoderm types (Smith *et al.*, 1988) whereas FGFs can induce all but notochord (Godsave *et al.*, 1988). It is therefore possible that the dorsal vegetal signal is TGFß-like whereas the ventral vegetal signal is FGF-like (Dale & Slack, 1987). Initially MIFs were isolated from heterogenous sources, so it was necessary to establish if there are similar factors present in the embryo at the correct time and in the correct distribution, before conclusions could be drawn about endogenous factors.

In *Xenopus* embryos there is a maternal transcript, *Vg1*, that is restricted to the vegetal hemisphere of the egg (Rebagliati *et al.*, 1985), and it has been shown to code for a factor related to TGFß (Weeks & Melton, 1987). This relationship suggests that *Vg1* is a good candidate for an endogenous mesoderm inducer. However, it does not qualify as a dorsal determinant since the mRNA is uniformly distributed dorso-ventrally (Dale *et al.*, 1989). It still remains possible that the active protein, produced by cleavage (Tannahill & Melton, 1989), is restricted to the dorsal side. The status of FGF as a ventral determinant is similarly uncertain. A *Xenopus* basic FGF (bFGF)

protein has been detected in the embryo but it has not been localised (Kimelman *et al.*, 1988; Slack & Isaacs, 1989). Furthermore this FGF lacks a secretory signal peptide (Kimelman *et al.*, 1988) and so cannot be transported from the cell by any known mechanism. It is possible that bFGF has an autocrine effect on perspective mesoderm cells in the animal pole (Smith, 1989).

More recently, it was demonstrated that the heterologous TGFß-related factor XTC-MIF, is the *Xenopus* homologue of human Activin A (Smith *et al.*, 1990). Activin A was originally characterised as a factor causing the release of follicle-stimulating hormone from pituitary cells (Vale *et al.*, 1986; Ling *et al.*, 1986). Activins are dimers made up of two $ß_A$ chains or a $ß_A$ and $ß_B$ chain. Although a homodimer of $ß_B$ chains has not been isolated, a synthetic $ß_Bß_B$ protein (Activin B) has properties similar to Activin A (Mason *et al.*, 1989). Thonsen *et al.* (1990) cloned *Xenopus* $ß_A$ and $ß_B$ genes and showed that $ß_B$ is expressed in the early blastula whereas $ß_A$ is expressed in the late gastrula. This implies that *Xenopus* Activin B is responsible for early induction and axial patterning. Furthermore, Mitrani *et al.* (1990) cloned the chicken equivelent of $ß_B$ and showed that it is transcribed precisely when axial mesoderm is being induced. A factor has also been purified from the mouse which has MIF activity and is antigenically related to Activin A (Sokol *et al.*, 1990), but the mouse genes have not yet been cloned.

The discovery that protein growth factors can act as MIFs is exciting since it brings us closer to understanding the mechanism(s) involved in induction. However, more distribution studies as well as experiments that eliminate these growth factors from the embryo are required, before we can draw conclusions about their endogenous roles. It has been suggested that the endogenous activity in *Xenopus* may be silenced by simply injecting specific antisense oligonucleotides into the oocyte (Shuttleworth & Colman, 1988).

During the first six hours of vegetal induction of animal cells, protein synthesis is essential (Cascio & Gurdon, 1987). The terminal mesodermal markers (for example actin in the case of muscle) are not detected for approximately 9 hours, when induction is completed. This is a sufficient time lag for the transcription and translation of regulatory genes whose products could

directly activate the terminal markers. Good candidates for involvement in such a cascade mechinism are homeobox-containing genes. This possibility is supported by the isolation of *Mix-1* from animal cells following induction with XTC-MIF (Rosa, 1989). *Mix-1*, which contains a homeobox, is rapidly induced (<30 mins). Surprisingly it is not expressed in mesoderm but in prospective endoderm of the embryo. *Xhox 3* is another homeobox containing gene that is activated by MIF induction (Ruiz i Altaba & Melton, 1989b). *Xtwi*, the *Xenopus* homologue of a *Drosophila* gene involved in mesoderm formation (*twist*), is expressed in early mesoderm cells of the embryo at a time that is consistent with a response to mesodermal induction (Hopwood *et al.*, 1989). *Xtwi*, with a *myc* related DNA-binding domain, is therefore another candidate to directly respond to the mesodermal inductive signal and subsequently regulate downstream genes.

1.2.2 Pattern formation.

Pattern formation involves the spatial organisation of differentiating cells into the correct order for morphogenesis to take place. More specifically, it involves the organisation of cell interactions, such as those described in the previous section, to bring about the correct three-dimensional pattern of tissues. In addition to the systems of early induction, the systems of patterning and regeneration of the insect and vertebrate limb have been particularly valuable in the study of pattern formation.

In 1969, Wolpert formulated his ideas on the relationship between positional information and differentiation. Some aspects of the theoretical concepts are still valuable. A developmental field is conceived as a group of cells that have their position specified with respect to the same set of points. By direct measurement of developmental fields from a number of systems (Wolpert, 1969) it appears that most fields are smaller than 50 cells in a linear dimension. This implies that large organs may be organised in subdivisions (Maynard Smith, 1960); they may be organised early but lose their overall field properties as they grow. The phenomenon of pattern regulation, where altering the size of a field by removal or addition does not disturb the overall pattern, was

discovered by limb manipulations and reveals the ability of cells to reinterpret new relative positional values. Wolpert (1969) has suggested that there may be one universal mechanism by which positional information is obtained. While this is an attractive possibility it does not seem likely given the diversity found in experimental systems.

It has been suggested that cell surface molecules may be involved in determining the morphogenetic response of cells to inductive signals through altering the physical behaviour of the cells (Gallin *et al*, 1986). However, it seems likely that if such factors do play a role, they complement a complex system of specific morphogens.

A morphogen is the term given to a signalling substance involved in pattern formation and morphogenesis and may include the inducing factors described in the previous section. There are a number of different ways in which morphogenetic interactions could be organised and some of these are outlined below. Specific references are taken from systems of amphibian induction but the models may be extended to other systems.

(1). There may be different factors produced in different regions, each specifying different fates. This is the situation suggested by the three signal model of mesoderm induction. The model is supported by the differential effects of two classes of MIFs on cells in culture, but there is as yet no evidence from the localisation of factors.

(2). Responding cells may experience different concentrations of the same factor; a morphogenetic gradient (section 1.2.5). This is supported by the finding that exposure of cells to different concentrations of XTC-MIF leads to different mesodermal fates (Smith *et al.*, 1988). High concentrations induce notochord while progressively lower concentrations induce muscle, followed by mesenchyme and mesothelium. This would require a gradient of inducing factor across the field of induction. No evidence for this is seen in the distribution of *Vg1*. However morphogenetic gradients exist in other systems (section 1.3).

(3). A more elaborate timing mechanism was proposed by Smith (1989) which includes the production of secondary signals by the responding cells. This invisages the responding cells passing through phases of development during which they are capable of differentiating, first into notochord, then into muscle etc. This follows the sequence of cell differentiation in the embryo. For differentiation to occur there must be a threshold level of a secondary signal, the level of which depends on the proportion of cells responding. If the secondary signal builds up slowly then the cells pass through the time in which they have a capacity for dorsal fates and can only form ventral cell types. If it builds up rapidly then dorsal cells are determined. It should be noted that the model still requires graded differences in the type or concentration of inducing factor to which the cells are exposed. Support for the model comes from experiments showing differences in the timing of cell behaviour changes induced by different classes of MIFs (Cooke & Smith, 1989). Changes induced by XTC-MIF begin at the early gastrula stage whereas those induced by bFGF occur at the mid gastrula stage. The induction of a secondary factor is supported by the work of Gurdon (1988).

(4). Homeogenetic induction, which involves the induction of mesoderm by mesoderm for example, may be involved in the subdivision of tissue layers to different fates. This type of effect has been directly demonstrated by the induction of mesenchyme and mesothelium from ectoderm by previously induced notochord and muscle (Kurihara & Sasaki, 1981).

It is possible that a combination of the above mechanisms are in operation together or at different times in the same system. It is now important to determine the precise mechanisms that lead to the induction of specific regulatory genes, such as *myo D* in muscle cells and *Xtwi*, for example, in lateral plate mesoderm and notochord. In this way we might understand the hierarchical relationship between different steps involved in differentiation, and how they are coordinated in the context of a three-dimensional embryo.

1.2.3 Anterioposterior (AP) positional determination and neural induction.

The process of gastrulation, described in section 1.1, leads to the arrangement of germ layers within the embryo with scope for new cellular interactions. During this process the basic amphibian body plan is determined (Gerhart & Keller, 1986). Although there is some evidence to suggest that maternal effects influence the AP axis before gastrulation (Kao *et al.*, 1986), these effects appear to be labile (Gerhart *et al.*, 1984). AP fates are determined in the mesoderm as it gastrulates; invagination across the dorsal blastopore lip and the position reached within the embryo are both important. The effect of the dorsal blastopore lip was shown by the organiser graft of Speman (reviewed in Hamburger, 1988) and the effect of the extent of invagination was revealed by experiments that arrest migration (Gerhart *et al.*, 1984), leading to the production of structures appropriate to position. The importance of timing in the determination of AP fate was shown by organiser grafts performed at different stages during gastrulation: the later the graft was performed the less ability the cells of the dorsal blastopore lip had to induce anterior structures (Hamburger, 1988).

It has been suggested that protein growth factors (section 1.2.1) may also be involved in AP patterning of the embryo since the homeobox containing gene *Xhox3* is induced by these factors and occupies a graded AP distribution in the axial mesoderm of the embryo (Ruiz i Altaba & Melton, 1989a; 1989b). Furthermore, XTC-MIF induces low levels of *Xhox3*, whereas FGF induces high levels, suggesting perhaps that XTC-MIF and FGF induce anterior and posterior mesoderm respectively. To investigate the role of *Xhox3* in AP determination, large amounts of synthetic *Xhox3* mRNA were injected into fertilised eggs (Ruiz i Altaba & Melton, 1989a). Mesodermal movements were not affected in the treated embryos but anterior structures were suppressed. It is important to note that *Xhox3* inhibited anterior development but did not induce posterior development (ie. not a homeotic effect). *Xhox3* is therefore not sufficient for posterior development but it is involved in determining cell fate. Other homeobox genes are likely to be similarily involved.

In their new position following gastrulation, the mesoderm cells induce overlying ectoderm to form neural structures appropriate to their position. Anterior neural tissue differentiates into brain, while in posterior regions spinal cord is formed. This has been shown to be largely the effect of the nature and the timing of the mesodermal signal (Hamburger, 1988; Slack, 1983). Therefore the mesoderm cells are responsible for both neural induction and conferring AP polarity on the ectoderm.

Work by Sharpe *et al.* (1987) has shown that a homeobox-containing gene, *XlHbox6*, is induced specifically in posterior neural tissue by mesoderm. In addition to identifying an important regulatory link in posterior neural induction the work shows that the uninduced ectoderm may be predisposed to an anterior or a posterior fate since *XlHbox6* expression is most easily induced in posterior ectoderm. This once again illustrates the multilayered nature of embryonic interactions.

It has been suggested that the effect of mesoderm on overlying ectoderm is not simply linear but also involves a neuralisation factor released by Spemans' organiser (Hamburger, 1988). Sharpe & Gurdon (1990) have recently examined neural induction in *Xenopus* using valuable early neural markers. All of the results can be explained by a linear induction with no evidence of a neuralisation factor along the plane of the ectoderm.

In the mouse embryo, AP positional values in early neural tissue are also evident. At the very earliest stages the headfold defines presumptive brain from presumptive spinal cord. By $8^3/_4$ days the segmental units of the hindbrain are visible (Lumsden & Keynes, 1989) but, as is shown by the work presented in this thesis, at even earlier stages (from 8 days) there is restricted expression of homeobox-containing genes and a gene containing a zinc finger DNA binding motif within the presumptive hindbrain (Wilkinson *et al.*, 1989a; Murphy *et al.*, 1989; Wilkinson *et al.*, 1989b; Murphy & Hill, 1991). These genes demonstrate an early response to the initial factors defining AP positional values and appear to be closely related to the segmentation process that defines

precise domains within the hindbrain. As regulatory genes themselves, they are likely to be involved in interpreting this information. This will be discussed more fully in later sections.

1.2.4 The role of neural crest cells.

As previously described (section 1.1.3) neural crest cells migrate from the dorsal edge of the neural folds as they are closing and contribute to the patterning of the mesoderm. There is an important difference between neural crest cells in the trunk and those in the head. Exchanges of neural crest cells in different regions of the trunk lead to normal development (Le Douarin, 1982) indicating that the cells are not patterned in the neural plate but are patterned according to the positions to which they migrate. It is possible that interactions along the migration pathway are important. On the other hand transplantations of neural crest cells from the midbrain to the hindbrain led to the development of an ectopic beak in the neck of the recipient chicken (Noden, 1983). This implies that cranial neural crest cells are patterned before thay leave the midbrain or hindbrain. Therefore, in the terminology of Sidney Brenner, cranial neural crest cells are European (ancestry is of primary importance) and trunk neural crest cells are American (acquired position is of primary importance) (McKay, 1989). This also indicates differences in the systems of determination in the anterior and posterior developing CNS.

1.2.5 Morphogenetic gradients.

Positional values may be conferred by cells experiencing different concentrations of the same factor (morphogen) (as described in section 1.2.2). The idea of morphogenetic gradients is attractive since it is difficult to conceive the existence of enough factors to confer all positional values seperately. It is likely that there is a variety of factors and that at least some of these confer a range of information depending on their concentration. The possibility that endogenous protein growth factors operate through gradients has been discussed (section 1.2.2). There is good

evidence that a morphogenetic gradient is involved in the patterning of the vertebrate limb bud (section 1.2.6), however, we must go to the *Drosophila* system to find proven examples of gradients of signalling substances. The *dorsal* and *bicoid* genes are known to be involved respectively in dorsoventral and anterioposterior polarity in the *Drosophila* embryo and the proteins have now been shown to be distributed in appropriate gradients (Steward *et al.*, 1988; Driever & Nusslein-Volhard, 1988).

As early as 1937, theories were proposed to explain embryological phenomena involving gradients of substances to which cells respond in a discontinuous way, at 'thresholds' (Dalcq & Pasteels, 1937). More modern theories (Lewis *et al*, 1977) suggest that the morphogen is produced at a local source and is received and broken down by dispersed cells with discrete thresholds which determine how the cells respond (the model of 'local source, dispersed sink, discrete thresholds'). Such models can explain a number of features of pattern formation such as pattern regulation (Wolpert, 1969), organiser grafts (Hamburger, 1988) and the barrier effect, where a gap in the sequence of structures follows insertion of an impermeable barrier (Slack, 1987).

Following Wolperts' suggestion (1969) that developmental fields were likely to span less than 50-100 cells in a line, Crick (1969) calculated that a molecule with a diffusion constant of 10^{-6} cm^2 s^{-1} could diffuse across this distance within 10 hours (the time required for induction, Wolpert, 1969). This figure has been adjusted by Slack (1987) to 10^{-8} cm^2 s^{-1}. According to the work of Maestro (1984) this can be achieved by small molecules or extracellular proteins. It has been suggested that gap junctions may be involved in transporting morphogens. However, by using antibodies that block gap junctions it has been shown that they are not necessary for mesoderm induction (Warner & Gurdon, 1987).

The existence of threshold responses is an important part of the gradient model and is supported by the induction of different types of mesoderm by different concentrations of MIFs (Smith *et al.*, 1988). It is not known if cells respond individually or as a population to these threshold levels. In the first case an individual cell would have several thresholds whereas in the

second case a group of cells would differentiate together depending on the proportion of those cells that are induced. Experiments show that individual cells are not capable of responding to an inductive signal (Gurdon, 1988). Further support for a cell population effect is suggested by the work of Cooke *et al.*, (1987) who examined explants where the cells experience different concentrations of soluble inducing factor. Cells were found to differentiate together.

Theoretical models concerning the diffusion of morphogens across a developmental field are unlikely to precisely reflect the realistic situation. There may be a contribution from the responsive cells, through their ability to respond and perhaps produce secondary factors, in setting up functional gradients in real systems. However the models serve as a useful framework in which to test the possibilities.

1.2.6 Retinoic acid: A potential morphogen in the vertebrate embryo.

Retinoic acid (RA), a metabolite of vitamin A (retinol), is known to have dramatic effects on a number of different cell types and embryonic systems. It induces differentiation in teratocarcinoma cells, such as mouse F9 cells which are induced to form parietal endoderm (Strickland & Mahdavi, 1978), suggesting a role for RA very early in development. It is also essential for the differentiation and maintenance of epithelial cells both *in vivo* and *in vitro* (reviewed by Lotan, 1980; Brown *et al.*, 1985; Asselineau *et al.*, 1989).

Probably the most striking effect of RA, and that which originally led to the proposal of a morphogenetic function, was the respecification of the pattern of digits in the chick limb bud (Tickle *et al.*, 1982). It had previously been shown that transplantation of cells from the posterior margin of the limb bud, the zone of polarising activity (ZPA), to the anterior margin led to mirror image duplications of the digits (Tickle *et al.*, 1975). This suggested that the ZPA is the source of a morphogen which determines the pattern of digit formation by setting up a concentration gradient across the AP axis (Maden, 1985). Local application of variable concentrations of RA to the

anterior margin of the limb bud replicated the effect of ZPA transplantations and in addition the extent of the effect after both treatments was found to be dose dependent. The discovery that physiological levels of RA are present in the developing limb, with a higher concentration at the posterior margin (Thaller & Eichele, 1987), supported the theory that the endogenous morphogen released by the ZPA is RA. Recently a related compound, 3,4 didehydroretinoic acid (ddRA), which is 6 times more abundant in the limb bud, has also been shown to have the ability to induce AP duplications (Thaller & Eichele, 1990), indicating that the system may involve more than one active compound. The effect of RA on limb morphogenesis has also been demonstrated in the regenerating urodele limb (Maden, 1982).

When high levels of retinoids, including RA, are applied to whole developing vertebrate embryos the effects are extremely teratogenic and include craniofacial and limb malformations. The craniofacial effects, seen in the brain and the neural crest cell derivatives have been characterised in the rat (Morriss, 1972; Morriss & Thorogood, 1978; see section 5.1) and also in the human following inappropriate administration of a retinoid drug, isotretinoin (13 cis retinoic acid; Accutane) (Webster *et al.*, 1986; Lammer *et al.*, 1985). Most of the craniofacial malformations can be explained in terms of an early defect; the shortening of the neural plate at the anterior end (Morriss & Thorogood, 1978). This leads to a shift in the relative position of the germ layers in the head and subsequent skeletal abnormalities can therefore be explained by incorrect neural crest cell migration (section 5.1). In the *Xenopus* embryo, RA causes anterior truncations in a dose dependent manner in which anterior neural tissue is transformed to a more posterior specification (Durston *et al.*, 1989, Sive *et al.*, 1990). It therefore appears that in mammals and amphibians RA is involved in pattern formation along the AP axis of the head.

Limb malformations (Kochar, 1973; Satre & Kochar, 1989) are induced by exposure to RA at later times during embryonic development than craniofacial malformations. This indicates that RA may be involved in at least two different systems of pattern formation in the embryo. This is no surprise given that there are at least three different nuclear receptors for RA in mammals. These belong to the superfamily of steroid/thyroid receptors which are known to act as

specific modulators of gene expression when bound by their appropriate ligand (for reviews; Evans, 1988; Green & Chambon, 1988). This indicates the mechanism by which RA controls gene expression, as was observed in F9 teratocarcinoma cells (Colberg-Poley *et al.*, 1985; Mavilio *et al.*, 1988; LaRosa & Gudas, 1988b; Simeone *et al.*, 1990). The structure of steroid/thyroid receptors is highly conserved. A ligand-binding domain (E) binds the activating substance (e.g. hormone, RA etc.); a DNA-binding domain (C) binds to enhancer-like elements in the responsive genes; and a third region (A/B) seems to affect accessibility to these enhancers and is thought to be involved in determining enhancer specificity. The RA receptors include RARα (Petkovich *et al.*, 1987; Giguere *et al.*, 1987), RARβ (Brand *et al.*, 1988; Benbrook *et al.*, 1988) and RARγ (Zelent *et al.*, 1989; Krust *et al.*, 1989). All three RARs have very high binding affinities for RA and finger-swap experiments were used to show that they are activated by RA (Green & Chambon, 1987). These experiments involve replacing the E domain of a hormone receptor for which target genes are known, with the E domain of an RAR and demonstrating stimulation of a target gene with RA. There is evidence that RARβ binds RA with a ten fold greater affinity (Brand *et al.*, 1988). RARβ also appears to be expressed earliest in the developing embryo (Ruberte *et al.*, 1991), and is the only receptor that has increased levels following RA treatment of F9 cells (Zelent *et al.*, 1989). This may be involved in amplifying the RA response. Another related nuclear receptor (RXR) has recently been isolated (Manglesdorf *et al.*, 1990) but its specific ligand has not yet been identified. It is likely that this binds to a molecule related to RA, perhaps the one mentioned above, ddRA, treatment with which also leads to digit duplications in the limb (Thaller & Eichele, 1990). The diversity of nuclear receptors is further increased by the generation of different isoforms by alternative splicing of the transcripts (Giguere *et al.*, 1990; Kastner *et al.*, 1990). This array of receptor molecules, each perhaps with a different binding affinity for RA and/or controlling a different set of responsive genes, is capable of conveying a range of information based on the level of available RA, either in the same or in different systems. This is consistent with the variety of effects observed when different cells or systems are exposed to excess RA.

Consistent with the idea of different RA receptors conveying different information is the fact that they are differentially expressed in the developing embryo (Zelent et *al.,* 1989, Ruberte *et al.*, 1990; Dolle *et al.*, 1989b; Dolle *et al.*, 1990; Ruberte *et al.*, 1991). There does not appear to be a gradient of expression in any part of the embryo which would reflect a gradient in cellular responsiveness, so it is unlikely that they contribute to setting up a gradient of information. However they are expressed in the correct regions to accept and transmit such information to the nucleus.

In the embryo there are also cellular binding proteins which specifically bind RA (CRABP) and retinol (CRBP) (reviewed by Chytil & Ong, 1984; Brockes, 1990). There are two related forms of each (see Dolle *et al.*, 1990). The binding proteins appear to mediate the effect of the potent retinoids. Dencker *et al.*, (1990) have recently shown that radiolabelled RA, that accumulated in specific regions of the embryo, was predominantly bound by CRABP. The binding proteins are also differentially expressed in the developing embryo (Maden *et al.*, 1988; Maden *et al.*, 1989; Dencker *et al.*, 1990; Dolle *et al.*, 1989b; Dolle *et al.*, 1990; Ruberte *et al.*, 1991). Consequently models have been proposed which suggest a role for binding proteins in determining the amount of RA that is made available to the nuclear receptor system (Robertson, 1988). Maden *et al.* (1988) showed that in the developing chick limb the level of CRABP is graded from anterior to posterior at the distal tip. This gradient is the opposite to that found for RA itself (Thaller & Eichele, 1987) and it was therefore proposed that CRABP steepens the gradient of functional RA by sequestering it at the anterior end. This offers an explanation for the previously disconcerting finding that the level of endogenous RA is sufficient to saturate receptors throughout the limb. It has been noted that the regions of the embryo that are most sensitive to excess RA correspond to the regions that express the highest levels of CRABP (Dencker *et al.*, 1990; Dolle *et al.*, 1990; Ruberte *et al.*, 1991). It has therefore been suggested by Dolle *et al.* (1990) that these regions require a low level of active RA for correct development, with CRABP sequestering RA and preventing it from stimulating the receptor system. On the other hand they suggest that

regions expressing CRBP require a higher level of RA which is supplied by the metabolism of CRBP bound retinol.

Genes that respond to RA-activated receptors have not yet been identified. Since RA appears to convey positional information, probably through a functional gradient across a developmental field, it is likely that at least some of the target genes will themselves be regulatory genes that participate in conveying finer levels of positional information. In *Drosophila*, the morphogenetic gradients established in the early embryo are known to be involved in the hierarchical process of segmentation which is controlled in part by homeobox genes (section 1.3.5). In vertebrates, homeobox genes have been shown to be induced following RA treatment of teratocarcinoma cells (Colberg-Poley *et al.*, 1985; Mavilio *et al.*, 1988; LaRosa & Gudas, 1988b). Interestingly the genes of the human *HOX 2* cluster respond differentially to RA depending on their position within the cluster (Simeone *et al.*, 1990). The genes at the 3' end of the cluster, which are also the genes expressed more anteriorly along the AP axis, are induced by lower concentrations of RA (section 1.4). *Hox 1.6* and *Hox 2.9* are induced in mouse F9 cells following RA treatment (see section 3.2.5). Homeobox genes are therefore candidate responsive genes for RA.

1.3 Drosophila development and the role of the homeobox.

The *Drosophila* system is very suitable for genetic analysis with numerous advantages related to ease and speed of handling and breeding. Therefore isolation and analysis of mutations is relatively simple. The field of developmental genetics, among others, has benefited from this type of analysis and we currently know more about how the body plan is established in *Drosophila* than in any other organism. Through mutagenesis studies genes have been isolated that determine the DV axis (reviewed by Anderson, 1987) and the AP axis (reviewed by Akam, 1987). Here, the

patterning of the *Drosophila* AP axis through the process of segmentation, and the role played by homeobox-containing genes in this process, are described.

1.3.1 Drosophila segmentation

During development of the *Drosophila* embryo, it is divided into repeating units (segments) along the AP axis. These segments reflect the basic organisation of the body plan. Through extensive genetic analysis more than 50 genes that are involved in the progressive determination of segments have been identified (Lewis, 1978; Nusslein-Volhard & Wieschaus, 1980; Wieschaus *et al.*, 1984; Jurgens *et al.*, 1984; Nusslein-Volhard *et al.*, 1984; Perrimon *et al.*, 1984; Schupbach & Wieschaus, 1986; Akam, 1987). Many of the genes have now been cloned (Akam, 1987) and *in situ* hybridisation analysis of expression has complimented genetic evidence in revealing the functions and interactions of individual genes involved in this process.

Each *Drosophila* segment is further divided into an anterior and a posterior compartment, compartments being units of cell lineage restriction (Garcio-Bellido *et al.*, 1979). From morphological and expression studies on the early embryo it appears that the developmental unit of the fly is a parasegment rather than a segment (Martinez-Arias & Lawrence, 1985). Parasegments consist of the posterior compartment of one segment and the anterior compartment of the next more posterior segment.

The genes involved in segmentation are divided into groups according to the time at which they act and the type of information which they convey (see below). These regulatory genes interact with each other, both within and between groups, in defining precise spatial expression patterns. The interactive relationship between two genes can easily be investigated in the *Drosophila* system. Mutations in each of the genes are available and using cloned probes for *in situ* hybridisation, the effect of a single mutation on the expression of other genes can be observed. However it is impossible to say if the genes interact directly or indirectly by this method. It seems

that the interactions are both hierarchical (Howard & Ingham, 1986) and combinatorial (O' Farrell & Scott, 1986). Specific examples of the best characterised interactions are given in section 1.3.5.

The groups to which genes involved in segmentation are assigned are described below.

(1) Maternal effect genes: These are expressed by the mother during oogenesis and are responsible for early, relatively broad , positional determinants. Maternal determinants are only present at the poles of the egg before fertilisation. Three types of mutation reveal that there are determinants for the anterior end (e.g. *bicoid* (*bcd*); Frohnhofer & Nusslein-Volhard, 1986). the posterior end (e.g. *oskar*; Schupbach & Weichaus, 1986) and both ends simultaneously (e.g. *torso*; Schupbach & Weichaus, 1986). The determinative nature of the *bcd* gene was demonstrated by the partial rescue of *bcd* mutants on injection of anterior cytoplasm from a wild type egg (Frohnhofer & Nusslein-Volhard, 1986). *In situ* hybridisation showed the accumulation of *bcd* transcripts at the anterior pole of the egg as expected (Frigerio *et al*, 1986). After fertilisation the transcripts form a concentration gradient along the AP axis with the highest concentration at the anterior pole, demonstrating the mechanism by which *bcd* influences the expression of later segmentation genes (section 1.3.5). Maternal transcripts of the *caudal* (*cad*) gene also have a graded distribution, from posterior to anterior in this case (Macdonald & Struhl, 1986). However, it is unlikely that maternal *cad* has a critical role in determining position as mutant eggs fertilised by wild type sperm have no phenotypic effect. In other cases the distribution of transcripts, and the effect of mutations, indicate that maternal effect genes participate in the initial events that trigger position specific activation of zygotic gene expression.

(2) Segmentation genes: These are further subdivided into three classes; gap genes, which cause the deletion of multiple adjacent segments from the embryo, pair rule genes, which cause the deletion of alternate segments and segment polarity genes, which cause pattern defects in each of the segments. Among the earliest zygotically expressed genes are the gap genes including *hunchback* (*hb*), *Kruppel* (*Kr*) and *knirps* (*kni*) which are involved in progressively more posterior regions of the embryo. The domains of the gap genes, as defined by mutations, overlap (Nusslein-

Volhard & Weischaus, 1980), but the domains in which *hb* and *Kr* are expressed are smaller and appear to abut (Jackle *et al.*, 1986). This anomoly was explained following an analysis of expression patterns in gap mutants by Jackle *et al.* (1986) which showed that the gap genes interact in defining their mutual expression boundaries. In this way gap genes define the first precise spatial subdivisions of the embryo.

Gap genes also regulate the next set of genes to be transcribed, the pair rule genes. This is revealed by gaps and alterations in the spacing of pair rule expression stripes in gap mutants (Carroll & Scott, 1986). When pair rule genes such as *fushi tarazu* (*ftz*) and *hairy* (*h*) are first expressed they are detected in broad continuous domains along the AP axis of the embryo, but within about 30 minutes, the continuous domains break down into seven distinct stripes of expression each spanning and separated by two segments (Hafen *et al.*, 1984; Weir & Kornberg, 1985). So the pair rule genes first set up the segmental pattern at the time that the cellular blastoderm is forming (prior to this the embryo consists of a syncytium of nuclei in a common cytoplasm). Furthermore the dynamic expression patterns indirectly indicate that the pair rule genes regulate each other in setting up this pattern. More direct experiments reveal a hierarchy among pair rule genes; only *h*, *runt* and *even-skipped* (*eve*) are necessary for the normal pattern of *ftz* (Howard & Ingham, 1986; Carroll & Scott, 1986). The *ftz* protein itself is not necessary for setting up *ftz* stripes and so cannot be similarily important in generating the periodicity. *ftz* is more likely to be involved in stabilising the pattern (Akam, 1987). Pair rule gene expression continues to evolve during gastrulation and in some cases adopts a single segment periodicity, either by the splitting of existing stripes, as with *paired* (Kilcherr *et al.*, 1986), or by the addition of stripes, as with *eve* (Macdonald *et al.*, 1986).

The prototype segment polarity gene *engrailed* (*en*) first shows the final metamerism of the *Drosophila* embryo in its 15 stripe expression pattern (DiNardo *et al.*, 1985). Each stripe corresponds to the anterior boundary of a posterior compartment and genetic analysis reveals that *en* is required to maintain the lineage boundary that separates and defines the anterior and posterior compartments (Morata & Lawrence, 1975). *en* expression is initiated at the start of

gastrulation, about 40 minutes after the initiation of pair rule gene expression. Over a 30 minute period the pattern spreads from anterior to posterior. It is thought that the pair rule genes set up a pre-pattern to which *en* responds since mutations in any of these genes alters the pattern of *en* (Howard & Ingham, 1986; Akam, 1987). The relationship between *ftz* and *en* appears to be relatively simple since, in *ftz*- flies the *en* stripes that normally overlie a *ftz* stripe are absent (Akam, 1987). Other segment polarity genes are also expressed in single segment stripes; ie. *gooseberry* (*gsb*) (Bopp *et al.*, 1986; Cote *et al.*, 1987), and *wingless* (*wg*) (Baker *et al.*, 1987). *wg* transcripts are detected at the posterior boundary of each parasegment, just across a compartment boundary from *en*, so they may be both involved in defining the parasegment boundary (Lawrence, 1987).

(3). Homeotic genes: These specify segment identity and are responsible for the the production of correct body parts. This has been concluded from the phenotypic effect of homeotic mutations which transform one body part into another which belongs elsewhere in the fly (Lewis, 1978). The homeotic genes are contained within two gene clusters; the genes of the *Antennapedia* complex (ANT-C) control head and anterior thoracic segments (Kaufman *et al.*, 1980) and the genes of the *Bithorax* complex (BX-C, Lewis, 1978) control thoracic and abdominal segments. From extensive genetic analysis alone, Lewis (1978) proposed that each segment is specified by a unique combination of homeotic genes. Subsequently it has been shown that at the level of expression, each parasegment is defined by the combination and relative level of active homeotic genes, for example parasegment 6 is defined by high level expression of *Ultrabithorax* (*Ubx*) (Akam & Martinez-Arias, 1985; White & Wilcox, 1985), parasegment 4 by high level expression of *Antennapedia* (*Antp*) (Carroll *et al.*, 1986; Martinez-Arias, 1986) and more posterior parasegments by specific combinations of genes.

It may be limiting to think rigidly of homeotic genes as specifying segment identity, and the dividing line drawn between homeotic and segmentation genes may be rather artificial. It may be more useful to think of a single progressive process of positional determination with more restricted cell specification at later stages. The correct expression of homeotic genes is dependent on the expression of the preceeding segmentation genes, on the genes *polycomb* (Lewis, 1978) and

extra sex combs (Struhl, 1981; 1983) which repress homeotic genes in those segments where they are not required and on interractions between homeotic genes (see section 1.3.5). As development proceeds the metameric expression of homeotic genes breaks down as a cell specific pattern emerges (Struhl & White, 1985). This indicates additional functions for homeotic genes not directly related to segmentation, but at an even finer level of organisation.

1.3.2 Genes involved in Drosophila segmentation contain a homeobox.

The first homeotic genes were cloned before anything was known about their molecular nature. This was achieved through the techniques of chromosome "jumping" and "walking" (Bender *et al.*, 1983; Garber *et al.*, 1983) and by microdissection of chromosomes (Scott *et al.*, 1983). On subsequent analysis of the complex loci, ANT-C and BX-C, they were found to contain low copy number repetitive elements within the transcription units of the genes (McGinnis *et al.*, 1984). This 180bp element, termed the homeobox, was first localised within *ftz*, *Antp* and *Ubx* genes (Kuroiwa *et al.*, 1984; Scott & Weiner, 1984; Gehring, 1985; Regulski *et al.*, 1985). It has subsequently been found in other homeotic genes (Harding *et al.*, 1985), in at least 5 segmentation genes (*ftz*, *eve*, *prd*, *en* and *gsb* (Bopp *et al.*, 1986; Cote *et al.*, 1987)) and at least two maternal effect genes (*bcd*; Frigerio *et al.*, 1986; and *cad*; Mcdonald & Struhl, 1986). Gap genes are the single class of segmentation genes not represented by a homeobox containing member. Some of the segmentation genes contain homeoboxes that are more divergent from the prototype *Antp* sequence, for example *en* (Poole *et al.*, 1985) and *invected* (Coleman *et al.*, 1987) and the *paired* sequence (Frigerio *et al.*, 1986).

The homeobox sequence is not only conserved within *Drosophila* genes but has also been found in many higher eukaryotes from annelids to chordates (Muller *et al.*, 1984; McGinnis, 1985) including higher vertebrates (section 1.4.1). It was not originally detected in a range of lower organisms including nematodes. However the recent report of 'scores' of homeoboxes in nematodes (Burglin *et al.*, 1989) indicates that there may also be related sequences in lower

organisms. The level of homeobox conservation indicates the important role that this element plays in the basic organisation of multicellular animals. Among the predictions made by Lewis (1978) about the homeotic genes, was a suggestion that they descended from a single ancestral gene to organise a segmental plan, since each appeared to be involved predominantly in a single segment. Lewis can not have imagined how dramatically this would be demonstrated by the structure and organisation of the genes (section 1.4.2).

1.3.3 The homeobox codes for a DNA binding domain.

The amino acid sequence encoded by the homeobox (the homeodomain) is similar to the regions within yeast mating type regulatory proteins Mat a1 and Matα2 (Shepherd *et al.*, 1984; Laughon & Scott, 1984). The region of similarity, in the 3' part of the homeodomain, is capable of forming a helix-turn-helix structural motif similar to the DNA binding regions of prokaryotic regulatory proteins (Anderson *et al.*, 1981; McKay & Steitz, 1981; Matthews *et al.*, 1982; Sauer *et al.*, 1982; Weber *et al.*, 1982; Anderson *et al.*, 1987). It was therefore suggested that the homeodomain also functions as a sequence specific DNA binding element. Recently the task of determining the three dimensional solution structure of a homeodomain by nuclear magnetic resonance spectroscopy has been completed (Qian *et al.*, 1989). This confirmed that the *Antp* homeodomain forms four α-helices with the predicted turn between helices II and III. X-ray crystallographic studies of prokaryotic protein-DNA complexes revealed that the second α-helix (helix III in Antp) fits into the major groove of the DNA (Wharton & Ptashne, 1987; Otwinowski *et al.*, 1988). Since amino acids in this region directly contact groups exposed in the target DNA, it is termed the recognition helix. Functional analysis of the homeodomain regions of *bcd* and *prd* have demonstrated that helix III is also involved in sequence specific recognition of DNA, although the important amino acids are in different positions than in prokaryotic proteins (Hanes & Brent, 1989; Treisman *et al.*, 1989).

The *Antp* helix-turn-helix region shares 13 out of 29 amino acids with Matα2 (Laughon & Scott, 1984) which acts as a repressor by specifically binding DNA (Johnson & Herkowitz, 1985). In addition to structural homology a number of more direct observations are consistent with a sequence specific DNA binding function for the homeodomain. Firstly, antibody localisation studies have shown that all homeodomain proteins tested to date accumulate in the nucleus (for example, *en*, DiNardo *et al.*, 1985). Secondly, in vitro DNA binding has been demonstrated for *en* (Desplan *et al.*, 1985) and *eve* (Hoey & Levine, 1988) among others. The binding of *eve* to sequences upstream of *en* and the *eve* gene itself supports expression studies in mutant flies which indicate that *eve* is involved in regulation of *en* and autoregulation of its own expression (Macdonald *et al.*, 1986). Thirdly, transient assays have shown that homeodomain proteins can regulate transcription in a site dependent manner in co-transfected *Drosophila* culture cells (Jaynes & O'Farrell, 1988; Han *et al.*, 1989; Dearolf *et al.*, 1989) and *Drosophila* embryos (Struhl *et al.*, 1989; Driever *et al.*, 1989; section 1.3.5). Furthermore from cotransfection assays it has been shown that homeodomain proteins can act as either activators or repressors (Krasnow *et al.*, 1989). Finally, human proteins that were isolated because of their activity as transcription factors contain divergent homeodomains (Scheiderett *et al.*, 1988; Bodner *et al.*, 1988).

The discovery of the homeodomain gave the first exciting clue to how *Drosophila* development is regulated at the molecular level. Other genes involved in segmentation also contain conserved protein motifs that tell us something about their function. Zinc finger DNA binding motifs, similar to that found in *Xenopus* transcription factor IIIA (Miller *et al.*, 1985), are present in products of the gap genes *Kr* (Rosenberg *et al.*, 1986) and *hb* (Tautz *et al.*, 1987). Again these have an inferred function in transcriptional regulation. The paired box, of unknown function, is found in the *paired* gene in addition to the homeobox, and also in two other genes including *bcd* (Bopp *et al.*, 1986).

How does the activity of segmentation genes as transcription factors fit together with what we know from genetic evidence about the role of these genes in the embryo? Given that during the early stages of *Drosophila* development the embryo is a syncytium, it is not difficult to conceive

how DNA binding proteins could act as determinants by localised accumulation and diffusion to set up information gradients. By the time that cellularisation occurs the segmental pattern has already been laid down, as was revealed by segmental expression of segmentation genes (section 1.3.1). At later stages positional information would have to be transmitted across cell membranes, as is the case for higher organisms. But DNA binding proteins would still be important as second messengers to the nucleus. It is therefore interesting that the product of one of the later expressed segmentation genes, *wingless*, is related to transforming growth factor b (Baker, 1987) and is a potential cell signalling substance.

1.3.4 Specificity of homeodomain DNA binding.

As argued in the previous section homeodomain-containing proteins appear to function through sequence specific DNA binding. However the proposed DNA binding regions, the homeodomains, of many of these proteins are more than 90% identical and have been shown to bind to the same sequences *in vitro* (Hoey & Levine, 1988). Different homeodomain proteins can also regulate gene expression in transient assay systems through identical target sequences (Jaynes & O'Farrell, 1988; Han *et al.*, 1989; Winslow *et al.*, 1989). Some degree of specificity has been demonstrated *in vitro* but it is difficult to reconcile the low level of specificity of homeodomain-DNA binding observed in artifical situations with the very specific roles that homeodomain proteins play during *Drosophila* development.

In the yeast system, specific binding by Mat α2 is achieved through combinatorial interactions with other proteins and results in cell-type specific gene expression. In haploid 'a' cells, α2 binds together with Mcm1, which normally activates 'a' specific genes, to a 32bp operator site and represses 'a' gene activity (Kelleher *et al.*, 1988). This co-operative binding of Mcm1 and α2 is mediated through protein-protein interactions. On the other hand in diploid a/α cells the α2 protein binds to a different set of regulatory sequences that bear only limited resemblence to the 'a' gene operators (Goutte & Johnson, 1988). This second set of operators are upstream of haploid

specific genes and are recognised by α2 in combination with a1. It is not known how co-operation with a1 leads to recognition of the haploid operator, it may be achieved through a change in conformation of the protein. It is quite likely that co-operative interactions play an important part in specific binding of homeodomain proteins in the *Drosophila* system as well. Inducible ectopic expression of the *Antp* gene attatched to a heat shock promoter has been found to cause transformations in limited regions of the embryo at specific times only (Gibson & Gehring, 1988). This 'buffering' of the developmental system against homeotic mutations could be achieved through combinatorial action of regulatory genes.

The function of conserved regions outside the homeodomain in homeodomain proteins, including the paired box and the POU domain, are not known. These may be involved in forming protein-protein interractions that alter binding specificity. While Han *et al.* (1989) have shown that combinations of homeodomain proteins can synergistically activate or repress reporter gene activity there is no biochemical evidence for such protein-protein interactions. However it has been demonstrated that the POU domain of the human genes *oct 1* and *oct 2* is required for high affinity DNA binding (Muller-Immergluck *et al.*, 1990; Gerster *et al.*, 1990). Another possibility is that binding of the appropriate regulators is achieved through competition for the same sites since different homeodomain proteins are known to bind to the same sites with different affinities (Desplan *et al.*, 1988; Hoey & Levine, 1988). There is suggestive evidence for such a mechanism (Jaynes & O'Farrell, 1988) but definite evidence is currently lacking.

Some studies have been designed to directly address the question of DNA binding specificity, but the results to date are somewhat contradictory. These studies examine the effect of altering various parts of the binding protein on binding specificity. Using mutated homeodomains in *in vitro* systems, two seperate investigations have indicated that a single amino acid, residue 9 of the recognition helix, is critical in determining specificity. Hanes & Brent (1989) found that changing this amino acid switched the specificity of the *bcd* protein so that it no longer bound to sites upstream of *cad*, but recognised sites normally bound by *Ubx* and *ftz*. Treisman *et al.* (1989) showed that a change in the same amino acid was sufficient to change the specificity of *prd* to that

of either *bcd* or *ftz*. On the other hand a homeodomain swop experiment indicated that other regions were also important in sequence recognition. This involved the substitution of a *Ubx* homeodomain in the *deformed* gene (Kuziora & McGinnis, 1989). A change in specificity was subsequently observed with the chimeric protein eventhough no changes in the putative recognition helix were involved. A different assay system was used by Mann & Hogness (1989) where ability of deletion mutants to produce correct inducible transformations in *Drosophila* embryos was tested. They concluded that the region important for *Ubx* binding identity was contained in the carboxy-terminus of the homeodomain, where the recognition helix is. However a similar assay of *sex combs reduced* and *Antp* suggested that in addition to residues within the homeodomain, residues adjacent to both ends are also important (Gibson *et al.*, 1990). It therefore appears that while single amino acid changes within the recognition helix are sufficient to change binding specificity in *in vitro* systems that other regions, possibly even outside the helix-turn-helix motif, may play important roles *in vivo*.

1.3.5 Regulatory interactions of homeodomain proteins.

Homeobox-containing genes involved in *Drosophila* segmentation interact with each other (section 1.3.1) and presumably with sets of downstream structural genes that are necessary for the formation of correct body parts. At the moment the only genes known to respond to homeobox genes are other members of the regulatory hierarchy. As previously mentioned, the original evidence for interactions between these genes, involving analysis of expression in mutant embryos, gave no indication whether the interactions were direct or indirect. Expression assays in transgenic flies or following transfection of cultured cells, are capable of addressing this question. The following are some specific examples of gene activity in which a homeodomain protein participates, and for which interactions at the molecular level have been characterised in some detail.

The most valuable study of homeodomain interactions to date has concerned the activation of *hb* transcription by *bcd*. As previously described (section 1.3.1), a gradient of

maternally expressed *bcd* is established in the fertilised egg (Frigerio *et al.*, 1986). Manipulation of this gradient by Driever & Nusslein-Volhard (1988) showed that *bcd* has all the properties of a morphogen in determining the position of subsequent body parts. Later expression of the gap gene *hb* in the anterior part of the embryo, is *bcd* dependent (Tautz *et al.*, 1987; Tautz, 1988) and the promoter region of *hb* has been shown to contain three high and three low affinity binding sites for *bcd* (Driever & Nusslein-Volhard, 1989). Manipulation of the number of *bcd* binding sites in transgenic flies (Struhl *et al.*, 1989; Driever *et al.*, 1989) altered the level of *hb* expression as well as the position and sharpness of the *hb* boundary: Increasing the number of high affinity binding sites led to a higher level of *hb* and a more posterior boundary. In other words the sensitivity of the *hb* gene to the concentration of *bcd* protein was increased. The elucidation of how *hb* expression is regulated by the concentration of *bcd* shows how relatively smoothly graded information of one gene product can be translated into discrete domains of target gene expression, ie. how a morphogenetic gradient can operate at the molecular level.

Section 1.3.1 described how parasegments can be defined by the level and/or specific complement of homeotic genes that are expressed. It was generally observed from mutant analysis that homeotic gene expression is repressed by genes expressed in more posterior parts of the embryo (for example, Hafen *et al.*, 1984) and therefore that unique combinations are brought about, at least in part, by interactions between the genes. Using a transient expression assay in a cell culture system, Krasnow *et al.* (1989) have shown that *Ubx* proteins repress transcription from an *Antp* promoter fusion but activate expression of a *Ubx* promoter fusion. This shows that the same protein can act as an activator or repressor depending on the promoter to which it binds. *Antp* and *ftz* proteins, on the other hand, activate both promoters (Winslow *et al.*, 1989). Another homeotic gene product, the *Deformed* protein, has also been found to autoactivate *deformed* expression in transgenic flies (Kuziora & McGinnis, 1988).

The type of analysis described in this section is now beginning to yield an understanding of the chemical functions of specific homeodomain proteins. Understanding the molecular

mechanisms by which these central regulators influence gene expression gives us some idea of how complex developmental systems can be built on simple molecular decisions.

1.4 Mouse homeobox-containing genes.

1.4.1 Sequence homology between Drosophila and mouse homeoboxes has allowed access to developmental genes in the mouse.

The mouse provides a very valuable genetic system in which more than 1300 loci have been identified (Lyon & Searle, 1989). It has been subject to developmental analysis for many years and as a result mouse embryogenesis has been very well described at a morphological level (section 1.1). However, important developmental mutations, such as those studied in *Drosophila*, have not been identified. This is mainly because of breeding difficulties like small litter sizes and long generation intervals and also the relative inaccessibility of the embryos in utero. The genes responsible for the limited number of mouse developmental mutations that have been identified through genetics cannot easily be isolated, primarily because of a low density of genetic markers. Insertional mutagenesis has also had only limited success in identifying developmentally important genes (Woychik *et al.*, 1990). However because of the remarkably high level of sequence conservation of the homeobox, it has been possible to identify a large number of homeobox-containing genes in the mouse that appear to play important developmental roles, as do their counterparts in *Drosophila*.

Original Southern blot analysis with *Antp* homeobox probes indicated that there were at least 10 homologous genes in the mouse (Ruddle *et al.*, 1985). The number now seems closer to 40 (reviewed by Kessel & Gruss, 1990). 30 of these are designated *Antp*-like and arranged within four

multigene clusters (described below). The remainder have homeoboxes that are similar to other distinctive *Drosophila* sequences. Two mouse genes, *En-1* and *En-2*, show similarity to *Drosophila engrailed* (Joyner & Martin, 1987; Davidson *et al.*, 1988). The *Drosophila* muscle specific *msh* box has at least two mouse counterparts; *Hox 7* (Hill *et al.*, 1989; Robert *et al.*, 1989) and *Hox 8* (Monaghan *et al.*, in press; Davidson *et al.*, in preparation) Murine *cdx-1* is related to *Drosophila caudal* (Duprey *et al.*, 1988) and recently two genes related to *evenskipped* have been reported in the mouse; *Evx-1* and *Evx-2* (Bastian & Gruss, 1990). The *Drosophila paired* gene contains two types of conserved motifs; the homeobox and the paired box. Of the 8 paired homologues in the mouse three contain both a paired box and a homeobox (Dressler *et al.*, 1990; Nornes *et al.* 1990). This growing list demonstrates the diversity of homeobox-containing genes in the mouse.

Other potentially important genes in the mouse have been identified through the existance of conserved sequence motifs, which also indicate the function of the gene. Among these are oncogenes, growth factor genes, zinc finger encoding genes and helix-loop-helix genes.

1.4.2 Many of the homeobox containing genes are clustered in the genome.

There are four clusters of homeobox-containing genes in the mouse genome, each spanning more than 100kb. The *Hox 1* cluster on chromosome 6 has ten members; the *Hox 2* cluster on chromosome 11 has 9 members; the *Hox 3* cluster on chromosome 15 has 4 members; the *Hox 4* cluster (formerly *Hox 5*, Kessel & Gruss, 1990) on chromosome 2 has 7 members (Colberg-Poley *et al.*, 1985b; Duboule *et al.*, 1986; Baron *et al.*, 1987; Rubin *et al.*, 1987; Hart *et al.*, 1985; 1987; Graham *et al.*, 1988; Awgulewitsch *et al.*, 1986; Breier *et al.*, 1988; Sharpe *et al.*, 1988; Featherstone *et al.*, 1988; Duboule & Dolle, 1989; Kessel & Gruss, 1990). Since the direction of transcription is the same for all the genes within a cluster, the clusters are said to have a 3' end and a 5' end. Examination of the individual gene sequences revealed that the clusters could be aligned with each other so that proteins encoded by genes in equivalent positions show greatest similarity. It was therefore suggested that the clusters arose from a single ancestral cluster by large scale

duplication events (Hart *et al.*, 1987; Duboule *et al.*, 1989). *Drosophila* homeobox genes are also clustered (Harding *et al.*, 1985) and it was shown independently by Graham *et al.* (1989) and Duboule & Dolle (1989) that the *Drosophila* ANTP and BX complex genes can also be aligned with the mouse genes. This is more clearly the case for some of the genes (eg. *Abd B* and *Dfd*) than others and it is interesting that the genes that correspond best are those that are thought to be most ancient (Akam *et al.*, 1988; Graham *et al.*, 1989). It is possible that these represent the original members of the ancestral cluster. However, the idea of a single common ancestral cluster requires that the genes of the ANTP and BX complexes became seperated during *Drosophila* evolution. This is supported by the fact that homologues of the ANTP and BX members in the red flour beetle, *tribolium*, are contained within a single cluster (Beeman, 1987).

It appears that homeobox genes are also clustered in other vertebrates, although with the exception of the human (Boncinelli *et al.*, 1988; Acompora *et al.*, 1989) the clusters have not been well characterised. There is at least one cluster in the chicken (Wedden *et al.*, 1989) and there are two related clusters in *Xenopus* (Fritz *et al.*, 1989). In zebrafish there are two genes related to *Hox 2.1* indicating duplication, but it is not known if they are part of gene clusters (Njolstad *et al.*, 1988).

It was originally noted by Lewis (1978) that the map position of a mutation within the BX complex is reflected in the position along the body axis at which the mutation had its effect. Following the cloning of the genes this was also shown to be true for the expression domains (Harding *et al.*, 1985). The clustered mouse genes are also expressed in domains which reflect their relative position within the cluster (Gaunt *et al.*, 1988; Graham *et al.*, 1989; Duboule & Dolle, 1989). This pattern together with the remarkable conservation of the clustered arrangement, strongly indicates that the genomic organisation is vital to the functioning of the genes. Graham *et al.* (1989) feel that this could be accounted for by the presence of cis acting regulatory elements dispersed throughout the clusters, which work over long distances to regulate co-ordinate expression. This is supported by the work of Zakany *et al.* (1988) on the expression of *Hox 1.3* constructs in transgenic mice; although large amounts of flanking DNA were included in the constructs correct expression was not achieved, implying that long range sequences are required

for regulation. An alternative suggestion (Peifer *et al.*, 1987; Gaunt & Singh, 1990) involves a regulatory mechanism based on chromatin conformation influencing gene activity. According to this model only 3' genes would be in an open and accesible conformation in anterior regions with progressively more of the cluster opening up posteriorly. However, it is not known how such structural changes based on position could be brought about and the model has not been supported by experimental evidence.

1.4.3 The expression patterns of mouse homeo-box containing genes reveal a role in conferring positional information during development.

The accumulated data on mouse homeobox-containing genes reveal that they are all transcribed in restricted spatial domains at crucial stages of embryonic development (Jackson *et al.*, 1985; Awgulewitsch *et al.*, 1986; Gaunt *et al.*, 1986; Dony and Gruss, 1987; Fainsod *et al.*, 1987; Gaunt, 1987; Joyner & Martin, 1987; Krumlauf *et al.*, 1987; Toth *et al.*, 1987; Utset *et al.*, 1987; Wolgemuth *et al.*, 1987; Breier *et al.*, 1988; Davis *et al.*, 1988; Davidson *et al.*, 1988; Featherstone *et al.*, 1988; Fibi *et al.*, 1988; Holland & Hogan, 1988a; Le Mouellic *et al.*, 1988; Schughart *et al.*, 1988; Sharpe *et al.*, 1988; Bogarad *et al.*, 1989; Hill *et al.*, 1989; Robert *et al.*, 1989; Gaunt *et al.*, 1989; Murphy *et al.*, 1989; Wilkinson *et al.* 1989b; Dolle *et al.*, 1989a; Frohman *et al.*, 1990; Murphy & Hill, 1991; for reviews, Holland & Hogan, 1988b; Kessel & Gruss, 1990). Bearing in mind the nature of the genes and the roles of their counterparts in *Drosophila*, this implies that they are involved in positional determination. Furthermore, detailed analysis of the precise expression patterns has indicated the particular systems that the genes are involved in and have revealed levels of organisation within the embryo that were not previously known. The roles of individual genes within these systems can only be discovered through gene manipulations and the establishment of mutant animals (section 1.4.4). However, the expression data have indicated the regions that may be affected by mutation and will help in the interpretation of what will, no doubt, be complex mutant phenotypes.

The earliest detected expression of homeobox-containing genes in the mouse is at $7^1/_2$ days, before somitogenesis has commenced. Only a subset of the genes that have been analysed are expressed at this early stage (Davis *et al.*, 1988; Holland & Hogan, 1988b) but already differences can be seen in the domains of expression. For example, *Hox 1.5* (Gaunt, 1987) and *Hox 1.6* (Sundin *et al.*, 1990) are expressed in posterior ectoderm and mesoderm whereas *Hox 3.1* (Gaunt, 1988) and *Hox 1.1* (Mahon *et al.*, 1988) are expressed only in the allantois. This relationship is maintained in the later embryo as the former two genes remain more anteriorly expressed than the latter two. It is important to note that distinctive patterns of homeobox gene expression are detectable during gastrulation when, as has been described earlier (section 1.1.1), positional identity is thought to be first established within the embryo. However, it is not known what signals are responsible for setting up the spatially restricted expression. Some possible mechanisms are suggested by the the induced expression of *Xenopus* homeobox-containing genes following exposure to mesoderm inducing factors (section 1.2.1) and by the differential response of homeobox-containing genes to RA treatment in tissue culture cells (Simeone *et al.*, 1990; section 1.2.6)

By the end of gastrulation all known homeobox-containing genes are expressed. In general the transcripts occupy large overlapping domains in the CNS, the peripheral nervous system (PNS) and the segmented mesoderm. There is also expression in subsets of developing visceral organs such as the kidney, stomach and lung. Most attention has been focused on expression in the somites and the CNS in attempting to understand how the genes may impart positional information. In general, although the domains overlap, they have sharp, stable and distinctive anterior boundaries within both tissue layers (Gaunt *et al.*, 1988; Graham *et al.*, 1989). There are also distinctive posterior boundaries but these are less sharp and appear to alter during development (Holland & Hogan, 1988b). Positional information could therefore be conferred by unique subsets of these genes being active in blocks along the AP axis. Since segmentation is used to organise at least part of the vertebrate body, it was of interest to determine how the expression domains relate to segments. It was found that in the $12^1/_2$ day prevertebral column, not only are

homeobox genes expressed in particular subsets of prevertebrae, but there are also differences in the levels of transcripts along the AP axis (Gaunt *et al.*, 1988; Sharpe *et al.*, 1988; Toth *et al.*, 1987; Holland & Hogan, 1988a; Gaunt, 1988). Each segment (prevertebra) may therefore be defined by the subset of genes active and the relative abundance of transcripts. Only recently were expression domains in the CNS aligned with hindbrain segments revealing a segmental pattern (Gaunt, 1988; Murphy *et al.*, 1989; Wilkinson *et al.*, 1989b; Chapter 3; Chapter 4).

The anterior boundary of expression of mouse homeobox genes in the mesoderm is generally more posterior than in the CNS (Utset *et al.*, 1987; Dony & Gruss, 1987; Toth *et al.*, 1987; Oliver *et al.*, 1988). This is not the case in *Xenopus* and it has been suggested that perhaps in the mouse, expression is in register in the two tissue layers at earlier stages but becomes more independent as development proceeds (Oliver *et al.*, 1988). However, this has not been demonstrated and in fact, does not appear to be the case for at least one mouse homeobox gene; *Hox 1.1* (Puschel *et al.*, 1990).

Analysis of homeobox gene expression in the limb has revealed how an individual developmental field within the embryo may be patterned. In the early *Xenopus* embryo the band of mesoderm from which the fore-limb bud will subsequently develop is distinguished by expression of *XlHbox1*, the homolog of *Hox 3.3* (Oliver *et al.*, 1988). Using *Xenopus* antibodies, it was shown that the protein has a graded distribution across the later developing limb bud in the mouse (Oliver *et al.*, 1988) with highest concentrations at the anterior proximal end. In contrast *Hox 4.4* (formerly *Hox 5.2*) is highest at the posterior distal end, indicating how positional identity within the limb may be established using information from a number of genes. In addition to these two genes other members of the Hox 4 cluster (Dolle *et al.*, 1989a) and the *msh* like genes; *Hox* 7 (Hill *et al.*, 1989: Robert *et al.*, 1989) and *Hox 8* (Davidson *et al.*, in preparation; Monaghan *et al.*, in press) are expressed with distinctive patterns in the developing limb. It has previously been described how position within a *Hox* cluster is reflected in position along the AP axis at which a gene is expressed (section 1.4.2). This relationship is also seen within the limb where 3' members of the Hox 4 cluster (formerly Hox 5) are expressed earlier and more proximal than 5' members

(Dolle *et al.*, 1989a). It appears therefore that some of the genes involved in positional determination along the AP axis are later involved in positional determination within the limb.

The expression of most homeobox-containing genes persists into later embryonic development, although the overall level seems to decline (Holland & Hogan, 1988b, Monaghan *et al.*, in press; Dony & Gruss, 1987; LeMouellic *et al.*, 1988). Some are still detectable in the CNS of newborn (Awgulewitch *et al.*, 1986; Utset *et al.*, 1987; Bogarad *et al.*, 1989) and adult mice (Bogarad *et al.*, 1989). In the case of *Hox 2.5* expression along the DV axis of the spinal cord becomes restricted to the dorsal horn at $13^1/_2$ days (Bogarad *et al.*, 1989). This fact, together with the detection of transcripts in the adult, indicates that the gene may have different roles at different times; it may be involved initially in determining position along the AP axis and later it may have a cell specific function within the CNS. Some *Drosophila* homeobox-containing genes are known to be similarily multifunctional (Struhl & White, 1985).

Presumably the information required for correct spatially restricted expression of homeobox-containing genes resides in adjoining control regions, although these may be widely dispersed (section 1.4.2). Attempts have been made to define these regions by examining the expression patterns of reporter genes attached to various amounts of flanking DNA in transgenic animals. It appears that, as in *Drosophila*, different controlling sequences are responsible for different elements of the pattern. Puschel *et al.* (1990) have examined the regions flanking *Hox 1.1.* They placed the *E. coli* Lac Z gene under the control of 3.6kb of DNA upstream, and 1.7kb downstream, of *Hox 1.1.* In embryos carrying this construct, they found that some elements of the normal *Hox 1.1* expression pattern were reproduced but others were lacking. The initiation of expression at $7^1/_2$ days and the establishment of the pattern over the next 24 hours were normal, as was the subsequent anterior boundary of expression in the neuroectoderm and mesoderm. However, no posterior boundary of expression was set and the transgene was expressed in mesodermal derivitives in which the endogenous gene was not. This implies that essential sites for some negative regulators were not present.

In another study, an altered *Hox 1.4* transcript was used as a marker for expression from transgenes containing 10kb and 5kb of DNA upstream of *Hox 1.4* (Wolgemouth *et al.*, 1989). These were both found to be transcribed in the correct temporal and spatial pattern but at a different level to the endogenous gene. This however may be related to the number and position of integration sites, or an altered stability of the manipulated transgene, rather than to a lack of sufficient controlling information.

As suggested previously, controlling sequences for clustered homeobox genes may be widely dispersed (Graham *et al.*, 1989; section 1.4.2). For this reason the analysis of a homeobox containing gene that is not part of a cluster of similar genes, may be more informative in the short term. Preliminary analysis of a lac Z fusion to 4.5kb of DNA flanking the *Hox 7* gene suggests that all the information necessary for faithful transcription is present (R.E. Hill, unpublished results).

1.4.4 Functional analysis of vertebrate homeobox-containing genes.

A number of lines of evidence, including the expression patterns described in the previous section, indicate that vertebrate homeobox-containing genes have a role in determining the body plan. Because of the high level of similarity between *Drosophila* and vertebrate homeodomains (section 1.4.1) and the more limited similarity with some known vertebrate transcription factors (Scheiderett *et al.*, 1988; Bodner *et al.*, 1988), vertebrate homeobox-containing genes are assumed to produce sequence specific DNA binding regulatory proteins. This is supported by the nuclear localisation of the protein products (Kessel *et al.*, 1987; Schulze *et al.*, 1987; Odenwald *et al.*, 1987) and the *in vitro* DNA binding activity of *Hox 1.5* (Fainsod *et al.*, 1986). However, specific roles for individual homeobox-containing genes have not yet been assigned.

Before homeobox-containing genes were precisely mapped, it was hoped that they would be allelic to known developmental mutations (Hill *et al.*, 1987; Holland & Hogan, 1988b). The mouse gene, *Pax-1*, contains a paired box similar to that within the *Drosophila* segmentation gene

paired (Deutsch *et al.*, 1988) but it does not contain a homeobox. It has been shown to be responsible for the developmental mutation *undulated* by the detection of sequence differences within the gene in three independent *undulated* alleles (Balling *et al.*, 1988; Kessel & Gruss, 1990). Undulated mutants have malformed intervertebral discs and this corresponds to the major site of expression of *Pax-1*. So it is now possible to analyse the specific developmental process in which *Pax-1* is involved. However, existing mutations have not been similarily assigned to homeobox-containing genes. The only remaining close linkage is between *En-1* and *Dominant hemimelia* (Hill *et al.*, 1987; Joyner & Martin, 1987). But these also have been shown to be genetically separate loci (Lyon & Searle, 1989; M. Higgins, personal communication). It is possible however, that there is a functional relationship between the adjacent genes.

It now appears that new mutations will have to be created in order to analyse the effect of homeobox-containing genes. Two categories of mutations can be produced; 'loss of function' and 'gain of function'. It is difficult to produce loss of function mutations in mammals, but procedures have been worked out and applied to a number of genes including HPRT, ß2 microglobulin and the proto-oncogenes *c-abl* and *int-1* (Kuehn *et al.*, 1987; Hooper *et al.*, 1987; Thompson *et al.*, 1989;Schwartzberg *et al.*, 1989; Zijlstra *et al.*, 1989 Koller *et al.*, 1989; McMahon & Bradley, 1990; Thomas & Capecchi, 1990). The principal is that the mutation is produced in embryonic stem (ES) cells in culture, by exchanging an altered form of the gene for the normal endogenous counterpart through homologous recombination. The ES cells can then be introduced into blastocysts to form chimeric animals. Homozygous mutants can subsequently be bred from the animals that transmit the altered gene through the germ-line. The most interesting outcome to date was the inactivation of *int-1* (McMahon & Bradley, 1990; Thomas & Capecchi, 1990). Like the homeobox-containing genes, *int-1* is homologous to a *Drosophila* segmentation gene; *wingless*. The mutant animals lack a large portion of the brain, including the midbrain and the anterior part of the himdbrain. This is the region in which *int-1* is expressed early in development ($8^1/_2$ days, Wilkinson *et al.*, 1987) and indicates that the gene is involved in the determination or subsequent development of this region. The production of 'loss of function' mutations in homeobox genes by

homologous recombination is currently an area of active research. No mutant animals have yet been reported but there has been success in inactivating *En-2* and *Hox 1.1* in ES cells (Joyner *et al.*, 1989; Zimmer & Gruss, 1989).

A loss of function mutation has been produced in a *Xenopus* homeobox gene by an alternative method. Wright *et al.* (1989) injected antibodies specific for the protein product of *XlHbox1* into the fertilised egg. These are presumed to bind to the endogenous protein and inactivate it. They found that in treated embryos the anterior neural tube, where *XlHbox1* is normally expressed, was transformed to a more anterior character. Here it resembled the hindbrain with a thin dorsal roof plate and a cavity similar to the 4th ventrical. This could therefore be interpreted as a homeotic mutation where spinal chord is transformed to hindbrain. The experiment however, took advantage of the ease of manipulation of the *Xenopus* embryo and would not be possible in mammals.

Gain of function mutations are more easily produced in higher vertebrates. The genes can be over-expressed, either in their normal domains or ectopically, after random insertion into the genome. The manipulated genes are injected into the fertilised egg where they are incorporated randomly to produce transgenic animals. The problem with this type of experiment is that the results, in general, are difficult to interpret. Wolgemuth *et al.* (1989) overexpressed *Hox 1.4* in this way. The transgene transcripts were stabilised by the presence of SV40 sequences at the 3' end, this also made them distinguishable from the endogenous transcripts. Since the transgene was driven by its own promoter, the transcripts were present in the normal *Hox 1.4* domains but at elevated levels. The phenotypic effect, seen in homozygous animals, was a condition called megacolon which is caused by improper innervation of the colon leading to insufficient peristaltic activity. There are two possible explainations for how overexpression of *Hox 1.4* could interfere with development in this way. *Hox 1.4* is normally expressed in the gut mesenchyme and overexpression here may be responsible for the inappropriate signals. Alternatively, *Hox 1.4* overexpression may interfere with neural crest cell behaviour since neural crest cells contribute to the ganglia of the enteric plexus which are responsible for peristaltic behaviour.

Balling *et al.* (1989) examined the effect of ectopic expression of *Hox 1.1* in the embryo. They used the ß-*actin* promoter to drive *Hox 1.1* expression. This led to almost ubiquitous expression in homozygous transgenic embryos and death of newborn mice within 14 days. The phenotypic effects were observed in the head, where neural crest cell derivitives developed abnormally, and in the vertebral column, where the atlas and axis were more like posterior vertebrae and an extra vertebra was found (Kessel *et al.*, 1990). Both mutant defects are seen in regions anterior to the normal domains of *Hox 1.1* expression and it is difficult to tell at present whether the effects are telling us anything about the specific role of *Hox 1.1* in the embryo or if they are non-specific effects of expressing a control gene in regions where it is not normally active. When more clear-cut data are available, like the effect of loss of function mutations for example, it will be easier to interpret this information. What this type of experiment does reveal however, is the capacity of homeobox-containing genes to alter normal developmental pathways when incorrectly expressed.

1.5 Brief outline of thesis.

The materials and methods used in the work presented here are described in chapter 2. In chapter 3, the characterisation of a previously unidentified mouse homeobox-containing gene, *Hox 2.9*, is presented. Also in chapter 3, several transcripts from the related gene; *Hox 1.6*, which were isolated from an $8^1/_2$ day embryonic cDNA library, are analysed to reveal which alternative splicing products are present in the developing embryo. Chapter 4 presents a detailed analysis of the expression patterns of *Hox 2.9* and *Hox 1.6* during development. Particular attention is paid to early stages (8-9 days) when hindbrain segments are being defined. In chapter 4 also, a comparison is made of the distribution of alternative splicing products of *Hox 1.6* in the embryo. The effect of excess RA on the developing hindbrain is examined in chapter 5. This is achieved by observing the expression patterns of the segmentally expressed genes; *Hox 2.9* and *Krox 20*, in

treated embryos. Chapter 6 contains a summary of the conclusions drawn from the preceding chapters and some suggestions for further work.

Chapter 2

Materials and Methods

Unless otherwise stated, enzymes were supplied by Boehringer Mannheim Gmbh, and other chemicals were supplied by BDH Ltd. In general, basic protocols outlined below are detailed in Sambrook *et al* (1989).

2.1 Phage library screening.

Libraries used:

(1) 8.5 day embryonic mouse cDNA library constructed in lambda gt10 (Fahrner *et al*, 1987).

(2) Mouse genomic library. The inserts were prepared by *Sau* IIIA partial digestion of genomic DNA and were cloned into the *Bam* HI site of lambda 2001 (received from T. Rabbits, ICRF).

Libraries were plated on LB-agar medium in 20 x 20cm plates at a maximum density of 1 x 10^5 plaques per plate (i.e. almost confluent lysis). To achieve this, *E. coli* of the appropriate strain, previously grown to stationary phase in LB-broth + 10mM $MgCl_2$ + 0.2% Maltose, were spun down and resuspended in 0.5 volumes of 10mM $MgCl_2$. 1ml of this bacterial suspension was innoculated with an appropriate volume of phage stock (previously titred) in a 50ml sterile tube and incubated at room temperature for 15 mins. The cells were then diluted with 40ml of molten Top LB-agar (40°C), swirled to mix, poured over a dried LB-Agar plate and incubated at 37°C overnight. The phage plaques on the resulting plates were overlaid with dry Nitrocellulose (Schleicher & Schuell) or Hybond-N (Amersham) filters which were keyed in place with a needle containing waterproof ink. Serial 'lifts' were taken from the same plate to facilitate multiple screening with the same or different probes. The first filter was left on the plate for 1 min., subsequent filters for 2 or 3 mins. The filters were processed by laying them on 3MM (Whatman) filter paper in 20 x 20 cm dishes saturated in the following solutions:

(1) 0.5M NaOH, 1.5M NaCl. 3 mins. (to denature the DNA).

(2) 1.5M Tris HCl, 1.5M NaCl (pH 5.5). 3 mins. (to neutralise the pH).

(3) 2 x SSC 3 mins. (to wash the filters).

The phage DNA was fixed to the dry filters by baking (Nitrocellulose: 2hr 80°C) or UV irradiation (Hybond-N: 1200mJoules). The filters were hybridised and washed as described for Southern blots (section 2.6.4) and were exposed to autoradiographic film for variable times (1-3 days at -70°C). Duplicate filters were always screened with the same probe, and plaques on the original agar plate which corresponded to duplicate spots on the film were selected as potentially positive (+ve) clones. This was done by cutting out the relevant area of top agar and placing it in 1ml of SM (50mM Tris, 5mM $MgCl_2$) in an eppendorf tube. This stock, which contained a mixture of +ve and -ve clones, could be stored at 4°C for several months while secondary and tertiary screens were carried out. This was done by repeating the procedure above, plating from the selected stock on 9cm petri dishes at a lower density until a pure stock of the +ve phage clones was selected. DNA was then prepared from the clone as described in section 2.4.3.

2.2 Subcloning of DNA fragments.

2.2.1 Plasmid and M13 cloning vectors.

pUC 9: One of the pUC series of plasmid vectors based on pBR322 (Bolivar *et al.*, 1977) retaining both the capacity for a high copy number and the ampicillin resistance gene (ß-lactamase) (Vieira & Messing, 1982). It contains an artificial polylinker inserted into the ß-galactosidase gene (*lac-Z*) providing a convenient test for recombinant plasmids: Non-recombinant plasmids are able to synthesise the enzyme which breaks down X-Gal to release a blue product. However in recombinant plasmids the *lac-Z* gene is interrupted by foreign DNA and the colonies remain white.

pTZ18U / 19U: The pTZ series of vectors was derived from the the pUC plasmid (Mead et *al*, 1986) and are 2.9kb. pTZ contains the F1 origin of replication and therefore can be prepared as single stranded molecules which are convenient for sequencing.

Bluescribe (Stratagene): A 3kb plasmid vector derived from pUC 19 (Messing, 1983). It contains all of the useful features listed for pUC 9 above and also contains the phage T3 and T7 promoter sequences flanking the polylinker enabling production of specific single stranded RNA transcripts, both sense and antisense.

Coliphage M13 vectors: M13 is a single stranded (ss) DNA filamentous phage of *E. coli*. Replication occurs via a double stranded (ds) replicative form (RF) giving rise to ss progeny virions. For this reason M13 has been developed as a cloning vector for DNA sequencing (Messing, 1983; Norrander *et al*., 1983). Foreign DNA is cloned into the polylinker of the ds M13 RF. Strains have been produced with the polylinker in different orientations so that clones can be sequenced from both ends. The *lacZ* gene has also been utilised to provide the blue/white test for recombinant molecules.

2.2.2 Bacterial strains.

All strains were *E. coli* K12.

JM83: Δ*ara*, (*lac-pro* AB), *rsp*L, F80, *lac* Z M15, (rk^+, mk^+) (Vieira & Messing, 1982). This was the host used for pUC based plasmids. The *lac* Z M15 gene is integrated into the host chromosome.

JM101: Δ(*lac -pro* AB), thi, supE, {ΔF'*tra* D36, *pro*AB, *lac* I^qZ M15} (Yanisch-Peron *et al,* 1985). This strain was used for the propagation of M13. JM101 contains the lac I^q mutation which overproduces the lac repressor so IPTG (isopropyl-ß-D-thio-galactopyranoside) must be added to induce ß-galactosidase synthesis (Muller-Hill *et al.,* 1968).

2.2.3 Bacterial media.

LB (Luria-Bertoni) broth and agar; Per litre: 10g bacto-tryptone (Difco), 5g bacto-yeast extract (Difco), 10g NaCl. pH 7.5. In addition LB agar contained 15g agar (Difco) per litre and LB top agar contained 6.5g agarose (Sigma) per litre.

H agar: Per litre; 10g bacto-tryptone (Difco), 8g NaCl, 12g agar (Difco), pH7.3. H agar plates were used to accomodate the *lac*Z blue/white test described above. The substrate for the test was also added to the plates; 0.02% BCIG (5-bromo-4-chloro-3-indolyl-ß-D-galactoside; stock solution = 2% in dimethylformamide). When using the JM101 vector, IPTG was also added at 0.024%.

Minimal agar: Per litre; 15g agar (Difco), 1mM $MgSO_4$, 0.1mM $CaCl_2$, 1mM thiamine-HCl, 0.2% glucose.

2 x TY broth: Per litre; 16g bacto-tryptone (Difco), 10g yeast extract (Difco), 5g NaCl. pH 7.0

Terrific broth (Tartoff and Hobbs, 1987): per litre; 15g bacto-tryptone (Difco), 30g bacto-yeast extract (Difco), 5ml glycerol, 1/10 volumes of 1MK_2HPO_4 added before innoculation. A richer medium than LB broth, it was used for large scale cultures for plasmid preparation since it gave a higher density culture and a better yield of DNA.

Media additives: Antibiotics to select for plasmid maintainance were used at the following concentrations: Ampicillin (Sigma), 100ug/ml; Tetracyclin (Sigma), 25ug/ml; Kanamycin (Sigma), 30ug/ml. Stocks were prepared at 1000x concentration and stored at -20°C. For addition of BCIG and IPTG see H agar above.

All media were sterilised by autoclaving. Additives were filter sterilised. Bacteria were grown at 37°C with good aeration for liquid cultures. Bacterial stocks were routinely made in 20% glycerol and frozen at -70°C.

2.2.4 Constructing subclones.

Subcloning DNA fragments into plasmid or M13 vectors for further analysis was routinely performed. DNA fragments of interest were either purified following seperation on a low melting point (LMP) agarose gel or 'shotgun' cloned into the vector to create random plasmid subclones. The subclone of interest could then be selected by analysis of miniprep DNA or by Grunstein-Hogness hybridisation (Grunstein *et al.*, 1975). The most consistently successful method used for purification of a DNA fragment was that of Burmeister & Lehrach (1989). The appropriate band was cut out of an LMP agarose (Seakem) gel visualised using long wave UV light. The agarose slice was melted at 65°C for 15mins. in 5mM EDTA pH8, 100mM NaCl. This was then equilabrated to 37°C and incubated overnight with Agarase enzyme (2 units / 100μl). The DNA was then cleaned by two serial extractions with hot phenol (68°C) (ie. by overlaying with an equal volume of phenol, mixing, centrifuging and removing the aqueous layer into a fresh eppendorf tube), and a chloroform extraction. The DNA could then be concentrated by ethanol (EtOH) precipitation (1/10 volumes of 3M NaOAc and 2.5 volumes of cold (-20°C) EtOH added, left at -20°C for at least 30 mins and centrifuged for at least 10 mins to pellet the DNA). The pellet was washed in 80% ethanol, dried under vacuum and resuspended in an appropriate volume of TE (10mM Tris-HCl, 1mM EDTA, pH7.5). The DNA of interest was then inserted into the vector using the enzyme DNA ligase prepared from phage T4. Ratios of vector to insert DNA were varied around a standard of 1:3 on a molar basis, to achieve maximum success. Sticky end (overlapping cohesive termini) ligations were more successful than blunt end ligations. In general ligations were carried out with 25ng of linear vector in a 10μl reaction volume (10x ligation buffer: 0.5M Tris-HCl pH7.4, 0.1M $MgCl_2$, 10mM Spermidine, 1mg/ml BSA, 10mM ATP, 0.1M DTT) with 1ul of T4 DNA ligase (~10 units) at 16°C overnight.

2.2.5 Amplification by polymerase chain reaction (PCR) and subcloning of the amplified region:

Synthetic oligonucleotides, complimentary to either end of the region of interest, were designed in the correct orientations to prime replication reactions into the required region in both directions. The oligonucleotides also contained noncomplimentary 5' tails with specific restriction enzyme sites which were ultimately used to directionally clone the amplified fragment. The principle of the amplification procedure is that *Taq*I polymerase, which is functional at high temperatures, is used to extend the primers. Multiple rounds of amplification are achieved by rapidly increasing the temperature to melt off the the newly synthesised fragments, cooling the reaction to reanneal the primers and repeating the replication reaction at 72°C.

Here, 100ng of both oligonucleotides and 10ng of linear template (cDNA insert in a phage vector) were placed in a 50μl reaction volume with 2.5mM dNTPs, 0.5μl *Taq*I polymerase in 1x PCR buffer. 30 reaction cycles were performed under the following conditions; 1.5 mins at 92°C to melt the DNA; 3 mins at 52°C to allow the primers to anneal; 3 mins at 72°C to allow DNA synthesis. 1/20 of the final solution was analysed on a mini agarose gel. When the required fragment was amplified successfully, it was restricted with the specific enzymes for which there were now sites at either end and ligated into the plasmid vector Bluescribe as described above (see figure 4.9).

2.2.6 Bacterial transformation.

E. coli JM83 cells were made competent for DNA uptake by the method of Simianis (Hanahan, 1985). The cells, with an approximate efficiency of 5 x 10^6 colonies per μg of supercoiled plasmid DNA, were stored at -70°C for later use. Cells were grown in LB medium plus 10mM $MgCl_2$ to mid log phase (A_{550} = 0.5). After chilling on ice and pelleting at 3krpm for 10 mins, cells were resuspended in 0.3 volumes of RF1 solution (RF1 = 100mM rubidium chloride, 45mM manganese chloride, 30mM potassium acetate, 10mM calcium chloride and 15% w/v

glycerol, pH 5.8) and left on ice for 30 mins. The cells were pelleted and resuspended in 0.1 volumes of RF2 solution (RF2 = 10mM MOPS, 10mM rubidium chloride, 75mM calcium chloride and 15% w/v glycerol, pH 6.8). After 15 mins. on ice, the cells were flash frozen in liquid nitrogen in 200μl aliquots and stored at -70°C. For transformations the cells were thawed on ice and the ligated plasmid DNA (<10μl) was added and left to adsorb for up to 1 hour. After a 90 second heatshock at 42°C, the cells were chilled on ice for 1 min, 0.8mls of LB medium added and the cells allowed to develop antibiotic resistance at 37°C for 1 hour. Suitable aliquots of the transformation were then plated on selective media and grown overnight at 37°C.

E. coli JM101 cells were used for transformation by M13 DNA. These cells need to be routinely streaked on glucose minimal medium plates, which select for cells carrying a plasmid encoding a gene involved in proline synthesis (the host cell has a deletion in this gene). The same plasmid is needed for synthesis of the F-pilus required for infection by M13. To prepare competent JM101 cells a culture was grown in 2xTY medium to A_{550} = 0.4. After harvesting, the cells were resuspended in 50ml cold 50mM $CaCl_2$ and left on ice for 20 mins. The cells were then repelleted and resuspended in 10ml of $CaCl_2$ solution, and stored at 4°C for up to 4 days. The transformation with ds M13 DNA was carried out as described for JM83 cells above. After heatshocking and cooling on ice, 200μl of a fresh exponential culture of JM101 cells was added and the mixture plated in H top agar on H agar plates (plus BCIG and IPTG as in section 2.2.3). Following growth overnight at 37°C recombinant phage containing plaques could be identified by the blue / white ß-galactosidase colour test (section 2.2.1).

2.3 Culture and treatment of murine F9 cells.

2.3.1. Maintenance of cells in culture

Cells were grown in Eagle's modified Dulbecco's or RPMI 1640 medium (Flow Laboratories), supplemented with 10% foetal calf serum (Gibco Bio-cult) which had been heat-

inactivated at 56°C for 30 mins. All cultures were grown at 37°C in 25, 80 or 175cm^2 plastic flasks (Nunc), in a 10% CO_2 atmosphere.

Monolayer cultures were split when confluent. Cells were detatched from the surface of the flasks by trypsinising. Cells were washed twice with ~3mls trypsin (0.5g trypsin, 1g glucose, 0.1g EDTA, 50ml saline D concentrate/litre; saline D concentrate consists of 0.24g Phenol Red, 160g NaCl, 8.0g KCl, O.9g Na2HPO4, 0.6g KH2PO4/litre). Cells were then detached in fresh media by a final incubation with trypsin.

0.5ml batches of cells, containing ~1 x 10^6 cells, were stored in freezing medium (10% DMSO, 90% foetal calf serum) in liquid nitrogen.

2.3.2. *Treatment of F9 cells with retinoic acid.*

F9 cells were exposed to 5 x 10^{-7} M retinoic acid. This was done when the cells were not quite confluent but covered about 75% of the flask surface. A stock solution of 10^{-4} M retinoic acid was made in EtOH and the appropriate amount added directly to the medium (ie. 100μl added to 50ml of media). After exposure for 24 or 48 hours the cells were harvested and used either to produce RNA or protein.

2.4 DNA preparations.

2.4.1 Plasmid DNA minipreparation.

This method was used to prepare small (~20μg of pUC based plasmids) relatively pure amounts of plasmid DNA rapidly. 15ml of LB broth (section 2.2.3), with appropriate antibiotic added, was innoculated with a selected single colony. This was grown overnight at 37°C in a 50ml sterile tube. The cells were pelleted by centrifugation at 3krpm for 10 mins and the pellet

resuspended in 1ml of TELT buffer (TELT = 50mM Tris-HCl pH7.5, 62.5mM EDTA, 2.5mM LiCl, 0.4% Triton X100) with 1mg of lysozyme (Sigma). After transfer to eppindorf tubes the samples were boiled for 1 min, placed on ice for 7 mins and spun in a microfuge for 15-30 mins to pellet the *E. coli* debris and denatured chromosomal DNA. Between 0.5 and 0.8 mls of supernatant were removed, and the plasmid DNA precipitated with 0.6 volumes of Isopropanol. After spinning for 15 mins, the plasmid pellet was washed with 80% ethanol, dried and resuspended in ~20µl TE. The resulting DNA was sufficiently pure for restriction enzyme analysis but generally not for sequencing.

2.4.2 Large scale plasmid preparation.

500ml of LB broth or terrific broth plus appropriate antibiotic, was innoculated with a selected bacterial colony and shaken at 37°C overnight. The culture was decanted into two 250ml Sorvall centrifuge buckets and spun at 6krpm for 5 mins. The pellets were resuspended in 20ml of cold GTE (GTE = 1% glucose, 25mM Tris-HCl pH 7.5, 50mM EDTA), plus 10mg/ml lysozyme. After 15 mins on ice, 40ml of alkaline SDS (alkaline SDS = 0.25N NaOH, 1% SDS) was stirred in, to denature chromosomal DNA. After a further 10 mins, 30ml of high salt solution (3M potassium acetate pH4.5 with glacial acetic acid) was added to precipitate *E. coli* debris and chromosomal DNA. This was spun out at 12krpm for 30 mins. The supernatant was filtered through muslin to remove the course precipitate and 0.6 volumes of isopropanol added. Samples were left for 5 mins at room temperature and centrifuged at 10 krpm for 20 mins to pellet the plasmid DNA. Pellets were washed with 80% EtOH, lyophilised until almost dry and redissolved in 22ml TE. The plasmid DNA was then purified by CsCl gradient centrifugation. 24g CsCl (BCL) and 2ml ethidium bromide (10mg/ml, Sigma) were added and the tubes were left in the dark at room temperature for 30 mins to allow precipitation of any remaining bulk protein. This was then removed by spinning at 3krpm for 5 mins and the optical density of the solution was adjusted to between .930 and .945 by addition of CsCl. The samples were then centrifuged at 40krpm

overnight in a Sorvall Ultracentrifuge (20°C). Plasmid bands could be visualised over UV light and were removed with a fine tipped pastette. The ethidium bromide was removed by serial extractions with butanol until the samples were clear. DNA was then precipitated by addition of 2.5 volumes of 80% EtOH, DNA spooled when possible, washed in 80% EtOH, dried and resuspended in TE.

2.4.3 Bacteriophage DNA preparation.

Plating bacteria of a suitable strain were grown in LB broth + 10mM $MgCl_2$ + 0.1% maltose overnight. These were spun down at 3krpm and resuspended in 0.5 volumes of 10mM $MgCl_2$. About 5×10^5 phage were added to 1ml of cells and incubated at room temperature for 15 mins. Molten top-agarose (40 mls at 42°C) was added to the bacteria and plated onto a 20 x 20cm pre-poured LB-Agarose petri dish. This was incubated at 37°C overnight, without inversion. The plate was inspected for complete lysis of bacteria. 20 mls of SM (section 2.2.1) was added to the plate, which was then agitated at room temperature for 5 hours. The SM was then pipetted into a 50ml tube and spun at 3krpm for 20mins to pellet agarose aggregates. The supernatant was carefully decanted and 0.1 mg/ml RNAse and DNAse were added and incubated at 25°C for 30 mins. An equal volume of 20% PEG 6000 (Polyethylene glycol, MW 6000, Sigma), 2M NaCl was added to the tube to precipitate phage particles on ice for 2 hours. Prolonged precipitation times (overnight) increased final yields. Loose precipitates were decanted into 13ml tubes and spun at 10krpm for 20 mins to pellet the phage. The pellet was then resuspended in 4ml 50mM Tris-HCl pH7.5, 10mM $MgCl_2$ and extracted three times with chloroform to remove all traces of PEG. The phage particles were then disrupted by the addition of EDTA, pH8.0, to a final concentration of 20mM, followed by rapid extraction with phenol. Following a second phenol extraction and two chloroform extractions, the phage DNA was precipitated by addition of NaCl to 0.2M and two volumes of EtOH. The DNA pellet was resuspended in TE. This bacteriophage DNA was sufficiently pure for most enzymatic manipulations.

2.4.4 Oligonucleotide synthesis.

Oligonucleotides were synthesised and purified according to manufacturers instructions on an Applied Biosystems 381A Oligonucleotide Synthesiser by Doreen Chambers.

2.4.5 Single stranded DNA preparations.

The single stranded M13 molecules required for sequence analysis were prepared as follows: 100ml 2 x TY broth was innoculated with 1ml of an overnight culture of JM101 cells. An isolated recombinant M13 plaque was added to 1.5ml of these cells and grown at 37°C for 5 hours. The supernatant was respun to ensure no bacterial cells were remaining, then the phage particles were reprecipitated by the addition of 200μl of 20% PEG 6000 in 2.5 M NaCl. After 15 mins on ice the phage were collected by centrifugation and resuspended in 100μl TE. After phenol and chloroform extractions the ss M13 DNA was ethanol precipitated.

2.5 RNA Isolation.

2.5.1 Isolation of total RNA

This procedure was carried out as previously published (Hill *et al.*, 1985). Tissue culture cells, mouse embryos or adult organs from freshly killed mice were homogenised in 20ml of 8M Guanidinium Hydrochloride (BRL) in TE, for 45 seconds, using a ~~mortirised~~ motorised tissue mixer. The suspension was centrifuged for 10 mins at 8krpm and the supernatant carefully removed. Half a volume of cold EtOH was added and after storage at -20°C for 30 mins, RNA was pelleted by centrifugation at 10krpm for 15 mins. The pellet was resuspended in 15ml 6M Guanidinium Hydrochloride in TE and reprecipitated with 0.5 volumes of EtOH. After repelleting the 6M extraction was again repeated. The RNA pellet was now resuspended in 10mls of DEPC treated

H_2O (0.1% Diethylpyrocarbonate in double distilled water, autoclaved). Resuspension often required the use of a Dounce homogeniser. NaCl was added to 0.2M and the RNA was precipitated with 2.5 volumes of EtOH. This precipitation was repeated and the RNA was finally resuspended in 5mls of DEPC H_2O and stored at -70°C. RNA concentration was estimated from the absorbance of light at 260nm (RNA at a concentration of 40mg/ml has an $A_{260} = 1$).

2.5.2 Preparation of poly (A) + RNA.

This followed the method of Aviv and Leder (1972). Total RNA (1 to 2 mg) was diluted to 1mg/ml and heated to 65°C for 5 mins. An equal volume of 2x loading buffer (20mM Tris-HCl pH7.5, 0.5M NaCl, 1mM EDTA, 0.1% SDS) was added to the RNA. An oligo (dT) cellulose (Sigma) column (1ml volume) was poured and equilibrated with loading buffer. The RNA was then applied to the column and the eluate reheated and reapplied. The column was then washed with 5 volumes of loading buffer and then with buffer containing 0.1M NaCl. The flowthrough buffer was retained at this stage. Elution buffer (10mM Tris-HCl pH7.5, 1mM EDTA, 0.1% SDS) was then added (3 volumes) and 0.5 ml fractions collected. Each fraction was checked for RNA by spotting onto a dried agarose - ethidium bromide plate and visualising with UV. Positive fractions were pooled and precipitated with ethanol. The poly (A) + pellet was resuspended in DEPC H_2O at a concentration of 1μg/μl and stored at -70°C.

2.6 Southern and Northern blotting.

2.6.1 Genomic DNA digests.

Genomic DNA was stored at 4°C until required. 5μg of DNA per gel lane was digested with the appropriate restriction enzyme(s) under the conditions recommended by the suppliers. The standard DNA concentration for digestion was 0.1 to 0.2 mg/ml, with an enzyme excess of 3

units per microgram. Digestion was for a minimum of 4 hours for mammalian DNA, and was normally performed overnight. Reactions were terminated by addition of stop mix (30% Ficoll, 0.25% Orange G, 0.5M EDTA, and 10x TAE buffer) and were stored at room temperature prior to electrophoresis.

2.6.2. Phage and plasmid DNA digests.

Phage and plasmid DNAs were digested largely as above. Approximately 1μg of DNA was usually sufficient for analysis on a large (25 x 20cm) agarose gel. Complete digestion was usually achieved within 1 hour. For partial digestion, a time course experiment was carried out to determine the conditions required. 1μl of enzyme was added to approximately 6μg of DNA and 1/6 of the reaction was removed into stop mix after variable times (eg. 1min, 2min, 4min etc.). If the reaction occurred too quickly, the enzyme was diluted 1/10 in reaction buffer. Five time points at which partial digestion was achieved were further analysed by gel electrophoresis and southern blotting as described below.

2.6.3 Agarose gel electrophoresis of DNA fragments.

The concentration of agarose was varied depending on the size of DNA fragments to be resolved, for example, 1% gels were routinely used to seperate fragments below 2kb and above 200bp, whereas gels as low as 0.4% were used to separate very large fragments. For gels which were electrophoresed relatively slowly (large gels of 25 x 20cm at 40V overnight) TAE buffer was used (20 x TAE = 0.8M Tris-HCl, 0.4M NaOAc, 20mM EDTA, pH8.2 with glacial acetic acid). Mini gels (20ml volume) were generally electrophoresed rapidly (~80V) in TBE buffer as this buffer is less prone to overheating at higher voltages. (10 x TBE = 1M Tris-HCl, 0.8M Boric acid, 20mM EDTA, pH 8.3). The size markers routinely used were lambda DNA digested with *Hind*III and / or PhiX 174 DNA digested with *Hae*III.

2.6.4 Southern transfer.

This method of transferring DNA from an agarose gel to a nylon membrane broadly follows the method of Southern (1975). Following photography of the gel stained with ethidium bromide and visualised over long wave UV light, the DNA was denatured by immersion in 3 volumes of denaturing solution (0.5M NaOH, 1.5M NaCl) for 40 mins. This was followed by 30 mins in neutralising solution (1M Tris-HCl, 1.5M NaCl pH 5.5). The gel was then laid on a transfer apparatus; a wick of Whatman 17MM filter paper with a reservoir of 20 x SSC (20 x SSC = 3M NaCl, 0.3M Na citrate, pH7.4). A nitrocellulose (Schleicher & Schuell) or Hybond-N (Amersham) membrane was carefully placed in direct contact with the entire gel. Two sheets of Whatmann 3MM filter paper were placed over the filter followed by paper towels to a depth of ~5 cm. An even weight (~1 kg) was placed on the assembly and transfer carried out for at least 8 hours. The membrane was then baked for two hours at 80°C (nitrocellulose) or UV irradiated with 1200μ Joules (Hybond-N). For mini gel blotting, the denature and neutraliser steps were reduced to fifteen minutes and the transfer time could be reduced to a couple of hours.

2.6.5 RNA electrophoresis and Northern blotting.

RNA was electrophoresed in agarose under denaturing conditions to minimise secondary structure which alters mobility. Two types of denaturing gel were used, the second type listed below was found to be more reliable as the pH buffering was more stable.

(1) $NaPO_4$ gel: The agarose was melted in 10mM (final concentration) sodium phosphate buffer (1M $NaPO_4$ buffer pH6.5 = 70% 1M NaH_2PO_2 and 30% 1M Na_2HPO_2). After cooling to 55°C, 18ml of formaldehyde solution was added per 100ml and the gel poured quickly.

(2) MOPS gel : The agarose was melted in 1 x MOPS buffer (final concentration) (5 x MOPS = 0.2M Morpholinopropanesulphuric acid pH7, 50mM sodium acetate, 5mM EDTA pH8). Again after cooling, 5ml of formamide was added per 100ml and the gel poured.

Up to 10μg of total RNA or 2μg of polyA+ RNA was loaded per gel track. RNA was denatured for 10mins at 45°C in 2 volumes of 1.5x sample buffer (300μl deionised formamide, 108μl formaldehyde solution, 6μl 1M phosphate buffer). 0.1 volumes of loading mix (0.1% bromophenol blue in stop mix) was added to the RNA samples which were loaded onto the agarose gel and electrophoresed at 40V (15V overnight). The low voltage was necessary to prevent buffer failure which leads to RNA degradation. Buffer failure could be detected by a change in the bromophenol blue to a yellow colour. The gel was soaked in 1 litre of 2x SSC for 2 hours to remove the formaldehyde. Marker lanes were removed, stained with ethidium bromide and photographed. The remainder of the gel was blotted on a Northern transfer apparatus (as Southern with 10 x SSC in the reservoir) for at least 8 hours. Filters were processed as for Southern blots.

2.6.6 Hybridisation.

Pre-hybridisation and hybridisation of Southern and Northern filters were carried out in hybridisation mix that consisted of 5 x SSC, 0.1% BSA, 0.1% ficoll, 0.1% polyvinylpyrrolidine, 0.5% SDS, 0.5% sodium pyrophosphate and 100μg/ml denatured salmon sperm DNA. This solution was stable for several months at room temperature. Filters were sealed in plastic bags with about 10ml of hybridisation mix per 20cm^2 of filter. After prehybridisation at the appropriate temperature for at least 20 mins, the probe was carefully added to the bag and hybridised overnight (routinely). The bag was then carefully opened in wash solution of the appropriate stringency (standard wash was 2x SSC, 0.1% SDS, 0.1% sodium pyrophosphate) at the hybridisation temperature. The probe was disposed of and the filters were washed until the signal reached an acceptable level. The filters were then wrapped in clingfilm, marked with radioactive pen or

flourescent ink (Glojuice, IBI) and autoradiographed at -70°C with Kodak XAR-5 film. Films were developed in a Fuji RGB-2 Automatic Film Processor.

2.7 Preparation of radiolabelled nucleic acid probes.

2.7.1 Nick translation of DNA.

Nick translations were performed by the method of Rigby (1977). DNA (100ng to 300ng) was incubated with DNA polymerase I from *E. coli,* and bovine pancreatic DNAse I, in the presence of the four nucleotides, one of which was radioactively labelled. This allowed the nucleotides to be specifically incorporated into the DNA, as the polymerase removes and replaces tracts of DNA initiating at single stranded nicks produced by the DNAse. The reaction was carried out at 15°C which permitted template extension by Pol I, but prevented excessive DNA degradation by DNAse. Reactions were carried out in 10μl volumes with 1μl 10x salts (500mM Tris-HCl pH7.8, 50mM $MgCl_2$, 0.2mM dATP, dTTP, dGTP, and 100μg/ml BSA), 100ng DNA in 5μl ddH_2O, 3μl (=30μCi) ^{32}P dCTP and 1μl of DNA polymerase I / DNAse I mixture (0.4 units / μl) for 1.5 hours . Incorporation of radioactivity was assessed by measuring the specific activity following TCA (trichloroacetic-acid) precipitation of a 0.5μl sample. The radioactively labelled DNA was seperated from unincorporated nucleotides by passing it through a sephadex G50 column, which traps single nucleotides. The probe was denatured by boiling for 5 mins immediately before addition to the hybridisation bag.

2.7.2 Random priming of DNA.

Labelling by this method proved successful in cases where nick translation failed, presumably because the template DNA was not sufficiently pure. However, a weaker signal was generally achieved with random primed probes. The technique is based on the method of

Feinberg & Vogelstein (1983) where random hexanucleotides act as primers for polymerase chain extension in the presence of radiolabelled nucleotides. A BCL random priming kit was used. This was a convenient technique since as little as 10ng of DNA could be used and the reaction could be carried out in the presence of agarose which had been melted and diluted. Routinely 50ng of DNA was denatured in 10μl TE for 10 mins by heating to 100°C. The sample was chilled on ice, and 5μl of random prime buffer /dNTPs was added, followed by 4μl ^{32}P dCTP and 1μl Klenow fragment of DNA polymerase I. Random prime buffer consists of hexanucleotides in 4x Klenow buffer and 200μM dATP, dGTP, dTTP. The reaction was incubated at 37°C for 1-2 hours. Incorporation was assessed and the probe was isolated and added to the hybridisation in the same way as nick translated probes.

2.7.3 Oligonucleotide labelling.

Oligonucleotides were radiolabelled by T4 polynucleotide kinase. This enzyme transfers the γ-phosphate from ^{32}P γ-ATP to the 5' OH group of a dephosphorylated oligonucleotide. 30ng of oligonucleotide in 4μl TE, 1μl 10x kinase buffer (0.5M Tris-HCl pH7.5, 60mM $MgCl_2$, 50mM DTT), 3μl ^{32}P γ-ATP and 1μl T4 polynucleotide kinase were incubated at 37°C for 30 mins. Incorporation was checked by chromatography of a small sample on DEAE (Whatman) paper in 3M ammonium formate, unincorporated nucleotides move through the filter at the solvent front, whereas labelled oligonucleotides can be detected at the source. The reaction mixture was then directly added to the hybridisation bag.

2.7.4 Production of synthetic RNA for in situ *hybridisation.*

RNA probes were produced *in vitro* using the SP6, T3 or T7 phage polymerase systems (Kreig & Melton, 1984; Melton *et al.*, 1984). Briefly, a plasmid containing the insert of interest and a flanking SP6, T3 or T7 promoter sequence (eg. in Bluescribe or pTZ vectors) was linearised with

a restriction enzyme distal to the insert. After phenol and chloroform extractions and ethanol precipitation, the DNA was incubated with the appropriate polymerase and ribonucleotides, at 37°C for 20 mins. This was followed by addition of more fresh polymerase and incubation for a further 20 mins. Inclusion of a radiolabelled (^{35}S) nucleotide produced a single stranded RNA probe of high specific activity. The template was then removed by DNAse digestion and, following phenol and chloroform extractions, the probe was isolated by ethanol precipitation. A sample of the probe was then taken (1μl) and diluted in TE (20μl). This was divided between two glass fibre filters (Whatman), one of which was washed with cold TCA. The filters were placed in 'Aquasol' (Dupont) scintillation fluid and specific activity was estimated in a scintillation counter standardised for ^{35}S emissions. Transcripts larger than about 150bp cannot readily enter cells, and so for *in situ* hybridisation it was necessary to fragment these transcripts. Random fragments were produced by alkaline hydrolysis in 80mM $NaHCO_3$, 120mM Na_2CO_3 (pH10.2) at 60°C. The time of incubation varied, depending on the number of scissions required to reduce the transcripts to 150bp. This obviously depended on the size of the full transcript and was estimated using the following formula;

$$t(min) = Lo - Lf / 0.011Lf$$

where

Lo = transcript size (kb)

Lf = desired fragment length (ie 150bp)

0.011 = the number of scissions per minute

In initial experiments the process was followed by running samples, before and after digestion, on a denaturing agarose gel (section 2.6.5). The gel was then dried under vacuum and exposed to film for approximately 1-2 hours (figure 2.1).

Following digestion the transcripts were again recovered by ethanol precipitation and the specific activity estimated as above (a loss of ~20% was suffered by the digestion process). The probe was dissolved at a working concentration of 1.2 x 10^5 disintegrations per min/μl in hybridisation mix. Hybridisation mix = 50% formamide, 10% dextran sulphate, 1x Denhardts solution, 20mM Tris-HCl pH8, 0.3M NaCl, 5mM EDTA, 10mM sodium phosphate, 0.5mg/ml

yeast RNA, 50mM DTT. The probe was directly applied to tissue sections within 24 hours of preparation.

2.7.5 Probes used for in situ *hybridisation.*

The 3' *pst* I / *Eco* RI fragment of *Hox 2.9* (figure 3.8) was used in riboprobe preparations. For *Krox 20*, the probe was prepared from the 1.5kb *Apa* I / *Eco* RI fragment (Chavrier *et al.*, 1988). The *Hox 1.6* probe used was previously designated cDNA 1 (Baron *et al.*, 1987). This last probe includes the homeobox sequence, but the characteristic pattern of expression observed indicates that there is no cross-reactivity under the conditions used (section 2.8).

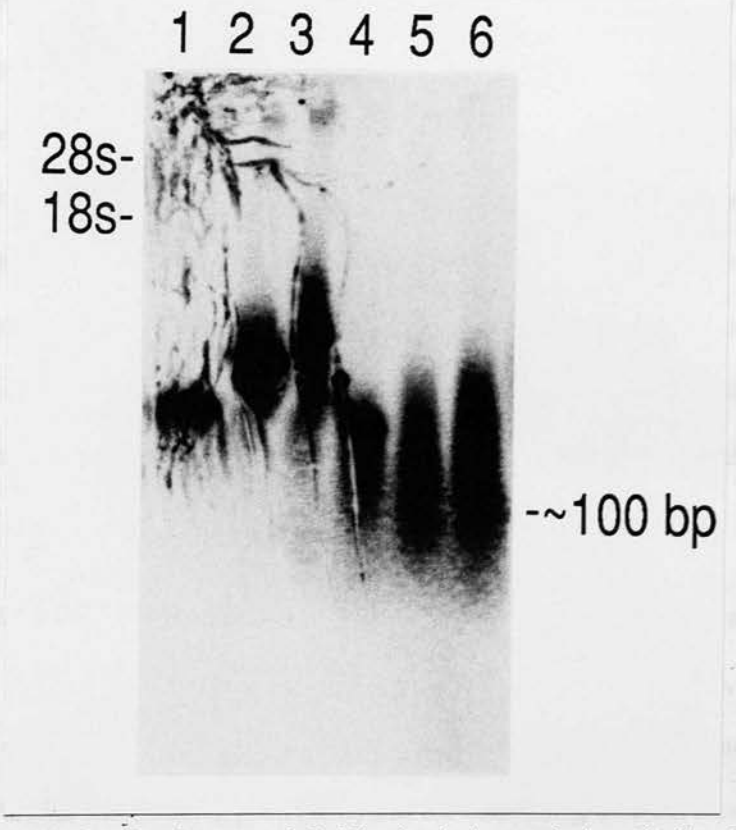


Figure 2.1. ^{35}S labelled probes for *in situ* hybridisation before and after alkaline digestion. Lanes 1 to 3 show undigested transcripts made from a 200bp, 400bp and 600bp insert respectively. Lanes 4-6 show the same transcripts after alkaline digestion for variable times, adjusted according to original transcript size using the formula given in section 2.7.4. The transcripts were separated on an RNA denaturing gel (section 2.6.5) which was then dried under vacuum and autoradiographed.

2.8 In situ hybridisation.

The procedures outlined below were developed by Duncan Davidson and Elizabeth Graham. They were previously published in part (Davidson *et al.*, 1988; Murphy & Hill, 1991).

2.8.1 Processing of embryonic material.

All embryos were obtained from outbred Swiss mice. These mice are albino and therefore present no problem with pigmentation obscuring the detection of label. Young female mice were placed in a cage with a proven male and on detection of a vaginal sperm plug the females were separated. The day of detection of the plug was designated day zero of pregnancy for nominal staging of the embryos. The embryos were more precisely staged by morphology after collection. Embryos were collected, following cervical dislocation of the mother, on designated days. They were dissected in ice cold phosphate buffered saline (PBS) and placed in 4% paraformaldehyde (PFA, pH7.2) dissolved in PBS, to fix, at 4°C. Embryos younger than 8 days were left in the desidual swellings, whereas older embryos were removed from extra-embryonic membranes, as much as possible. Embryos younger than 10 days were fixed for 4-5 hours, whereas those older were fixed overnight. Prolonged fixation makes the tissue brittle and difficult to handle.

After fixing, the embryos were dehydrated through a series of EtOH solutions in 0.85% NaCl; 50%, 2x 15 mins; 70%, 2x 30 mins; 85%, 30 mins; 95%, 60 mins; 100%, 3x 60 mins. The embryos were then cleared in toluene; 2x 30 mins, followed by overnight incubation at room temperature. Paraffin wax (melting point 56°C) was the embedding medium used for the embryos. Molten wax, at 60°C, was allowed to impregnate the tissues for 3 hours with 1 or 2 changes into fresh wax. Embedding was carried out in solid glass dishes that allowed easy manipulation and orientation of the embryos under a dissecting microscope. The wax was allowed to cool slowly at room temperature. The blocks could be stored indefinitely in a cool dry atmosphere.

5-7μm sections were cut through the wax embedded embryos on a standard microtome. These were floated on a 55°C waterbath and collected on glass slides that had previously been coated with TESPA (3-aminopropyltriethoxysilane, Sigma) to which the sections bind tightly. Sections were sealed onto the slides at 60°C overnight and were stored dessicated at room temperature.

2.8.2 Pretreatment of tissue sections.

Before hybridisation, the sections must be treated in a number of ways. Two 10 min washes in xylene dissolved the wax which was then removed in two serial 5 min washes in EtOH. The sections were then rehydrated through a series of EtOH solutions; 2 mins in each of 90%, 70%, 50% and 30%. Next, they were washed for 5 mins in 0.85% NaCl followed by PBS. They were then treated for 20 mins in 4% PFA in PBS pH7.2, and washed 2x 5 mins in PBS. A proteinase K treatment (20μg/ml in 50mM Tris, 5mM EDTA for 7.5 mins) was used to etch the surface of the tissue to facilitate access of the probe to the transcripts *in situ* in the cells. After another PBS wash, the sections were treated again in 4% PFA for 5 mins. Acetic anhydride treatment (625μl in 0.1M triethanolamine pH8, 2x 10mins) served to neutralise charged ions on the surface of the sections which may bind probe non-specifically. After final washes in PBS and 0.85% NaCl, the sections were again dehydrated by bringing them in the opposite direction through the series of EtOH solutions described above. They were allowed to air dry from 100% EtOH.

The treated sections were generally hybridised the same day but could be stored for at least four months if kept dessicated and dust free.

2.8.3 Hybridisation and post-hybridisation washing.

Radioactive RNA probes, dissolved in hybridisation mix (section 2.7.4), were directly applied to the dry tissue sections by pipette and overlaid with siliconised coverslips to avoid spreading and drying. The slides were hybridised at 55°C overnight in sealed slide boxes, in the presence of a tissue soaked in 50% formamide, 5x SSC to equilibrate the atmosphere. After hybridisation the slides were first washed for 20 mins in 5x SSC, 10mM DTT at 55°C to remove the cover slips and free probe. A 20 min high stringency wash followed at 65°C (50% formamide, 2x SSC, 0.1M DTT). This was followed by three 10 min washes in NTE (0.5M NaCl, 10mM Tris, 5mM EDTA, pH7.5) to remove all traces of DTT. The non specific hybridisation was reduced by a 30 min treatment with RNAse A to degrade probe that is not part of a stable duplex (40μg/ml RNAse A in NTE, pH 7.5). A second high stringency wash was then needed to remove the degraded probe (as above). The slides were finally washed 3x 10 mins in 2x SSC, 3x 10 mins in 0.1x SSC, dehydrated through an EtOH series as before and air dried.

2.8.4 Autoradiography of slides.

A 33% solution of Ilford K5 film emulsion was used to coat the hybridised slides. This was melted at 40°C in a waterbath (only under S902 safelight), mixed with two parts distilled water and placed in a dipping bath where the slides were dipped back to back. They were allowed to dry overnight at room temperature in a light tight box, after which they were packed away in sealed slide boxes, with silica gel, to expose at 4°C. The exposure time varied between 2 and 4 weeks with the particular probe used.

The slides were developed in Kodak D19 developer for 4 mins at 20°C and fixed in a 1/4 solution of Kodafix for 4 mins (also 20°C). After thorough washing in water, the slides were stained for approximately 30 seconds in 1% methyl green, allowed to air dry completely and mounted underneath coverslips in DPX mounting fluid. The slides were examined and photographed under high and low power magnification, using bright field illumination to visualise the tissue and dark field illumination to enhance the labelling.

2.9 DNA sequencing.

DNA sequencing was carried out by the dideoxy chain termination technique (Sanger *et al.*, 1977) using a T7 Sequencing kit ('Sequenase', Pharmacia). DNA fragments were prepared for sequencing in a number of ways (section 3.2.3) so that both double strand (ds) and single strand (ss) sequencing techniques were required.

2.9.1 Preparation of double strand DNA templates for sequencing.

A protocol was worked out for using ds DNA templates based on the recommendations of Pharmacia. Overall, ds DNA sequencing is less efficient than ss sequencing, however, it avoids the necessity for subcloning the fragment into a ss vector where it cannot be directly analysed by restriction enzymes in the ss form. Prior to the sequencing reaction the template must be denatured to allow access of the sequencing primer. This was achieved by NaOH denaturation. ~2μg of CsCl purified (section 2.4.2), supercolied plasmid DNA were added to 40μl of denaturation solution (0.2M NaOH, 2mM EDTA) and left at room temperature for 5 mins. 4μl of neutralising solution (0.4M NaOAc pH4.5) was then added, followed by 90μl of EtOH. The DNA was then spun for 4 mins and the pellet washed in 80% EtOH. After spinning again for 4 mins the pellet was dried. The pellet of denatured DNA could then be directly dissolved in the reaction mix for the first stage of the sequencing procedure. This procedure (outlined below) was largely the same for both ds and ss templates.

2.9.2 Sequencing protocol.

The first step of the sequencing reaction involved annealing the primer (either the universal or reverse M13 primers, or oligonucleotides specific to internal regions of the insert) to the template, followed by elongation from this primer by the enzyme T7 DNA polymerase. The

second step involved chain termination. For each template four separate sequencing reactions were performed. All contained primer, template and all four dNTPs (dATP was radiolabelled) but each contained a different dideoxynucleotide (ddNTP). Incorporation of a ddNTP into the newly synthesised DNA strand resulted in chain termination, as no 3'-OH group was available to form the next phosphodiester bond. A series of DNA fragments was therefore generated, all originating at the primer, and terminating at the sites of incorporation of the appropriate ddNTP.

The annealing reaction contained 2μg of template, primer at 0.8mM and annealing buffer ($MgCl_2$, DTT) in a total volume of 14μl. This was mixed and centrifuged briefly before incubation for 10 mins at 60°C. The samples were then allowed to cool slowly to below 37°C as the primers annealed. Previously denatured ds templates did not require such high temperatures and were usually allowed to cool from about 45°C, or were incubated at room temperature for 15 mins. 3μl of labelling mix (dGTP, dCTP and dTTP in solution), 10μCi a-^{35}S dATP (>400Ci/mM) in 1μl and 3 units of T7 DNA polymerase in 2μl of enzyme dilution buffer, were added and the reaction left for a further 5 mins. Four tubes labelled G, A, T and C containing 2.5μl of the appropriate termination mix were pre-incubated for 1 min before addition of 4.5μl of the labelling reaction to each tube. The components were mixed and incubated for 5 mins at 37°C. Reactions were then stopped by the addition of 5μl of stop solution (deionised formamide solution containing EDTA, xylene cyanol and bromophenol blue). The reactions could be analysed immediately on polyacrylamide gels or stored at -20°C.

2.9.3 Polyacrylamide urea gel electrophoresis of sequencing reactions.

The radiolabelled fragments produced by the sequencing reaction can be resolved on polyacrylamide gels (Sanger & Coulson, 1978) and detected by autoradiography. Sequencing gels were run on 50cm glass plates in TBE buffer (section 2.6.3). One glass plate was coated with MPTS (30μl g-methacryloxpropyltrimethoxy silane (Sigma) in 10ml EtOH + 300μl 10% glacial acetic acid). This treatment ensures adhesion of the gel to the plate. The second plate was silane

coated with 2% dimethyldochlorosilane in trichloroethane (BDH) to ensure that the gel adhered to the MPTS plate only. The plates were taped together seperated by 0.5mm PTFE spacers and gel mix added by pipette. The gel mix contained (per 200ml); 84g urea, 30ml 40% acrylamide stock (38% acrylamide, 2% bisacrylamide), 40ml 5x TBE and brought to 200ml with dH_2O. The crosslinking agent and catalyst were added immediately before pouring: 180μl 10% ammonium persulphate and 90μl TEMED (N'-tetramethylethylenediamine, BRL) per 40ml of gel mix.

After setting, the gel was put into a vertical gel apparatus, 1x TBE buffer added and the comb carefully removed. The wells were rinsed with buffer immediately to remove urea. The samples were heated at 80°C for 5 mins and 1.5μl of each added per well. Samples were electrophoresed at 27 watts for 90 - 100 mins for a standard run. The run was on occasion extended to as long as 3.5 hours to read further into the sequence. The gel was fixed in 10% methanol, 10% acetic acid for 10 mins and was then washed thoroughly in running water (>5 mins) and dried at 80°C. Overnight exposure to Kodak XAR-5 film was usually sufficient to clearly read 150-200 nucleotides, see figure 4.9.

2.9.4 Sequence analysis.

Alignment of sequences was carried out using the Amersham Staden Plus suite of programs on a DCS IBM AT clone. Sequence analysis and homology searches were carried out using the University of Wisconsin Genetics Computer Group's software package on the Daresbury computer.

2.10 Treatment of embryos with retinoic acid.

The following procedures were performed by Dr. Gillian Morriss-Kay at Oxford University.

C57BL/6 mice were housed 3 to a cage with males at 9am and checked for plugs at 11am, 12noon and 3pm. The time of observation of a plug was designated the start of day 0 of pregnancy. Retinoic acid (RA) was made up as 5mg RA in 0.8ml absolute ethanol to which 9.2ml arachis oil was added (0.5mg/ml solution). Crystalline RA was kept in the dark at 4°C for a maximum of 48hrs. Mice weighing approximately 25g were given 0.5ml (10mg/kg) or 0.6ml (12mg/kg) of this solution by oral gavage on day 7 + 18hrs or day 8 of pregnancy. Control dams of the same stages of pregnancy were given 0.5ml or 0.6ml of vehicle alone. Mice were sacrificed by cervical dislocation on day 9, day 9 + 12hrs, day 9 + 18hrs or day 10. Embryos were prepared for *in situ* hybridisation as described in section 2.8.1.

Chapter 3

Molecular characterisation of

Hox 1.6 and a closely related gene, Hox 2.9.

3.1 Introduction.

As described in section 1.4.2, the mouse *Antennapedia*-like homeobox genes reside within four tightly clustered multigene arrays in the mouse genome. The clusters appear to have arisen from a common ancestral cluster by chromosomal duplication events. This is shown by sequence comparison of the genes; for example, all but two genes in the *Hox 2* cluster have counterparts in the *Hox 1* cluster (Hart *et al.* 1987; Graham *et al.* 1989; Duboule & Dolle, 1989). The cognate genes within the clusters show a further level of similarity in that the genes are organised in the same linear order along the chromosome. It is also apparent that the mammalian clusters are remnants of an ancient ancestral cluster that pre-dates the organisation of homeobox-containing genes in insects. The insect complex of homeobox-containing genes, the HOM-C , which includes the *Bithorax* (Lewis, 1978; Sanchez-Herrero *et al.* 1985) and *Antennapedia* (Kaufman *et al.* 1980) gene complexes of *Drosophila*, is known to be homologous to the mammalian clusters from sequence comparison and from gene organisation along the chromosome. Closely related *Hox* genes, which share equivalent positions in different mammalian clusters, are therefore thought to have arisen from a common ancestor represented by a single *Drosophila* gene. On this basis the mouse genes are divided into subfamilies of paralogues.

A cDNA clone for *Hox 1.6*, the 6th member of the *Hox 1* cluster to be identified, was isolated in this laboratory (Baron *et al.* 1987). The homeobox sequence of *Hox 1.6* (Baron *et al.* 1987; LaRosa & Gudas, 1988b) revealed that it is most closely related to *Drosophila labial* (Mlodzik *et al.* 1988). *Hox 1.6* and *labial* are also similar in that they occupy the extreme 3' position in their respective clusters (Diederich *et al.* 1989; Baron *et al.* 1987). In an effort to identify other mouse *labial*-like genes, *Hox 1.6* was used to probe an $8^1/_2$ day embryonic cDNA library. A weakly hybridising cDNA clone was isolated in this way (Murphy & Hill, 1991). The characterisation of this clone, which represents a previously undescribed *labial*-like gene, is reported here.

Original analysis of *Hox 1.6* indicated that the gene is differentially transcribed (Baron *et al.*, 1987). It was therefore important to identify and characterise the alternative transcripts to determine the proteins which they encode and perhaps to understand the process by which they are produced. Two full-length alternative *Hox 1.6* transcripts were described by LaRosa & Gudas (1988b) following an investigation of transcripts induced by retinoic acid (RA) treatment of F9 teratocarcinoma cells. One of these transcripts encodes a homeodomain containing protein. The other transcript lacks a 203bp region 5' of the homeobox (figure 3.1) and can only encode a truncated protein without a homeodomain. In the study presented here, several embryonic *Hox 1.6* cDNA clones were isolated. These were analysed to determine the types of transcripts that are present, and their relative abundance, in the developing embryo.

In *Drosophila*, homeobox genes are regulators involved at all levels in the hierarchical process of segmentation (section 1.3.2). Also, as described in section 1.4.3, the spatially restricted expression patterns of vertebrate homeobox-containing genes are consistent with a role for these genes in spatial organisation of the embryo. However, in vertebrates the majority of homeobox genes were not found to be expressed in an obviously segmented manner (Holland & Hogan, 1988b) and we could only guess what underlying organisation they reflected. Vertebrate segmentation can be most clearly seen in the mesodermal somites (section 1.1.2) and in repetitive morphological structures in the developing hindbrain, called rhombomeres (section 1.1.4). Preliminary analysis of the expression of a second *labial*-like homeobox-containing gene is reported here and reveals a segmental pattern in the developing mouse hindbrain.

3.2 Results.

3.2.1 Isolation and mapping of Hox 1.6 *cDNA clones.*

A partial *Hox 1.6* cDNA clone, previously isolated by us (Baron *et al.*, 1987), was used to screen an $8^1/_2$ day embryonic cDNA library in lambda gt10 (Fahrner *et al.*, 1987). Twenty five cDNA clones, which hybridised to the *Hox 1.6* probe, were isolated in three successive library

screens. These clones were analysed and compared to the transcripts found in F9 teratocarcinoma cells by LaRosa and Gudas (1988b). The proportion of different types of transcripts present in the embryo could therefore be estimated.

DNA preparations from the 25 lambda clones were digested with *Eco* R1 to determine the size of the inserts (table 3.1). The digested DNAs were Southern blotted and the blots were hybridised with the *Hox 1.6* homeobox region. Two of the clones that hybridised weakly with *Hox 1.6*, hybridised more strongly with *Hox 7.1* (Hill *et al.*, 1989) indicating that they are not *Hox 1.6* clones. These, and six similar weakly hybridising clones were therefore eliminated from further analysis. The *Eco* RI fragments of the remaining 18 cDNA clones that hybridised to the *Hox 1.6* homeobox region must extend 5' of an *Eco* RI site 3' of the homeobox (see figure 3.1). Therefore, those fragments which hybridised with the homeobox and which are longer than 0.6kb extend far enough 5' to include the differentially spliced region. The particular *Hox 1.6* splicing products that these represented was determined by probing with a 16bp oligonucleotide complimentary to a sequence within the differentially spliced region (figure 3.2). Of the 17 *Hox 1.6* cDNA clones that could be categorised in this way, 10 contained the differentially spliced region and 7 lacked it (table 3.1). It is clear therefore that both transcripts are well represented in the $8^1/_2$ day embryonic library and are produced at $8^1/_2$ days of development.

The above results were confirmed by sequencing six of the cDNA clones (figure 3.1). It was found that three contained the differentially spliced region and that one had this region removed. Of the remaining two sequenced clones, one was truncated 3' of the alternately spliced region and so could not be categorised, while the other diverged from the published sequence within the hexapeptide region. The latter clone contained 130bp of new sequence 5' of the hexapeptide which was in-frame with the homeobox. This sequence was identified only once and it has not been established if it represents a third type of transcript or if it was produced as a cloning artifact.

Table 3.1 Analysis of *Hox 1.6* cDNA clones.

The columns denoted (*) list the fragments which hybridised with the Hox 1.6 homeobox (hbox) probe or the oligonucleotide specific for the differentially spliced region.

clone	insert size	Eco R1 fragments	fragments subcloned	subclones sequenced	*Hox1.6 h.box	*diff. spliced region	presence of diff. spliced region
cDNA 1	0.65kb	0.65kb	+	+	+	-	-
1^1	1.7kb	1.1kb	+	+	+		-
		0.6kb	+	+	-		
3^1	0.4kb	0.4kb	+	+	+		truncated
4^1	0.7kb	0.7kb	+	+	+		-
5^1	0.7kb	0.7kb	+	+	+		+
5^2	1.3kb	0.8kb			+	-	-
		0.5kb			-	-	
6^2	1.45kb	1.0kb			+	+	+
		0.45kb			-	-	
7^2	1.45kb	0.95kb			+	-	-
		0.50kb			-	-	
9^2	1.3kb	0.85kb			+	-	-
		0.45kb			-	-	
13^2	0.3kb	0.3kb			+		truncated
14^2	1.7kb	1.0kb	+		-	-	+
		0.7kb	+	+	+	+	
2^3	2.0kb	1.2kb			-	-	+
		0.8kb			+	+	
7^3	1.85kb	1.0kb			-	-	+
		0.85kb			+	+	
8^3	1.2kb	1.2kb	+	+	+	+	+
9^{3*}	1.0kb	1.0kb			-	-	-ve control
10^{3*}	1.0kb	0.7kb			-	-	-ve control
		0.3kb			-	-	
11^3	1.0kb	1.0kb			+	+	+
12^3	0.65kb	0.65kb			+	-	-
14^3	1.65kb	1.0kb			-	-	+
		0.65kb			+	+	
16^3	1.7kb	1.0kb			-	-	+
		0.7kb			+	+	
17^3	1.75kb	1.0kb			-	-	+
		0.75kb			+	+	

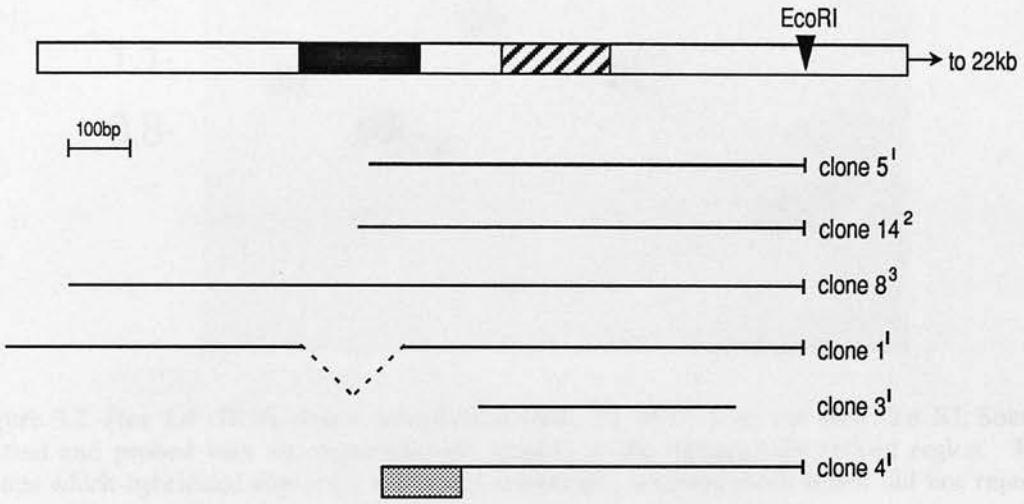


Figure 3.1 *Hox 1.6* cDNA clones which were sequenced to determine whether or not they contained the differentially spliced region (solid box). This region was absent from clone 1[1].

: homeobox; : previously undescribed sequence.

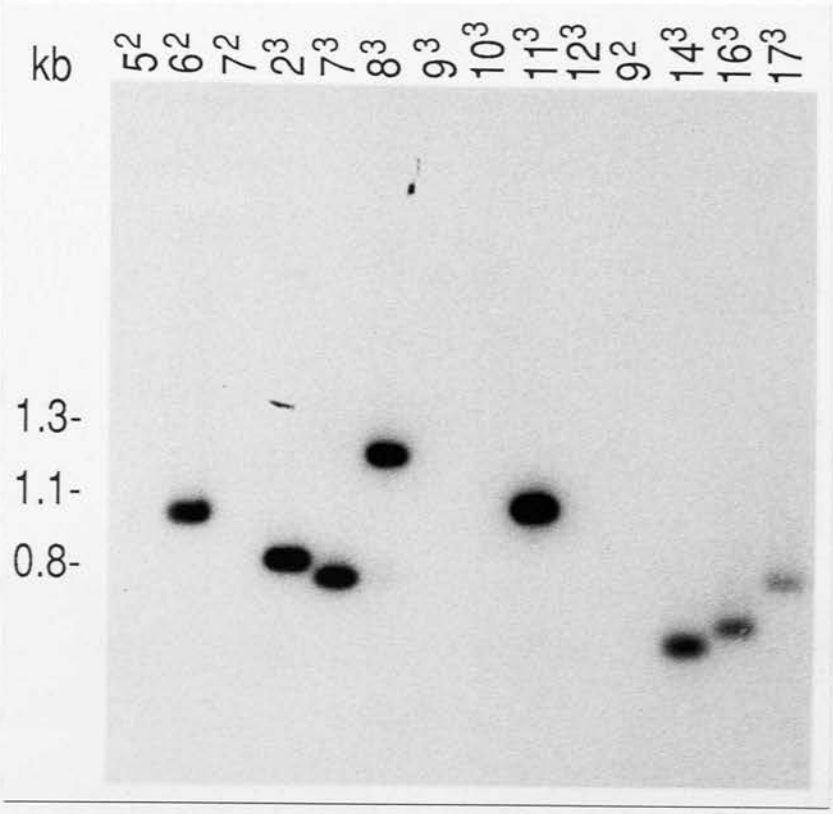


Figure 3.2. *Hox 1.6* cDNA clones described in table 3.1 which were cut with *Eco* RI, Southern blotted and probed with an oligonucleotide specific to the differentially spliced region. Those clones which hybridised represent full-length transcripts, whereas those which did not represent shorter spliced transcripts.

3.2.2 Isolation and chromosomal localisation of Hox 2.9*; a cDNA clone related to* Hox 1.6.

Hox 1.6 was used as a probe in a further screen of the $8^1/_2$ day embryonic cDNA library in which weakly hybridising clones were selected (Hill *et al.*, 1989). This was carried out in order to isolate previously unidentified homeobox-containing genes. Preliminary analysis of a 1.7kb cDNA clone (pL7) isolated in this way, revealed that it contains a homeobox sequence that is most similar to that of *Hox 1.6* (R. Hill, unpublished results). This was established by using a consensus homeobox oligonucleotide to prime a sequencing reaction across the homeobox.

The pL7 cDNA clone was subcloned in both orientations into the plasmid vector ptz18U, where it was mapped for the position of a number of restriction enzyme sites (figure 3.8). Several of the restriction fragments were subcloned into the Bluescribe vector and were tested for

specificity as hybridisation probes by screening Southern and Northern blots (some problems had previously been encountered by cross hybridisation to ribosomal RNA). The 3' *Pst* I / *Eco* RI and *Bam* HI / *Eco* RI fragments detected only one band on Southern blots and so were highly specific. The *Pst* I / *Eco* RI fragment (PRI) consistently produced a higher signal than the *Bam* HI / *Eco* RI fragment and so was used as a pL7 probe in all subsequent experiments.

An attempt was made to determine the chromosomal location of the gene that encodes the pL7 cDNA by screening panels of DNA from recombinant inbred mice (Bailey, 1971; Taylor, 1982). DNA from the progenitor strains, DBA and C57BL/6, were digested with 15 different restriction enzymes and screened for a polymorphism with the PRI probe. No polymorphisms were detected. However, this is not surprising given the high level of conservation of homeobox genes. To overcome this problem, intraspecific backcrosses the C57BL/6 inbred strain and *Mus spretus* (Robert *et al.*, 1985), were used to map the location of the gene. These species can be bred under laboratory conditions and, because of the evolutionary distance between them, they are polymorphic at many more loci than different strains of *M. m. domesticus*. DNA preparations from backcrossed offspring ([C57BL/6 X *Mus spr*]F_1 X C57BL/6) were obtained from G.-L. Guenet (Instituto Pasteur). These animals inherit one C57BL/6 chromosome and a second chromosome from either C57BL/6 or *Mus spretus* and can therefore be analysed for co-segregation of two markers. They are suitable for mapping purposes since they have previously been characterised for segregation of 67 polymorphic markers covering about 80% of the mouse genetic map (Robert *et al.*, 1985; 1989).

A pL7 polymorphism was detected between C57BL/6 and *Mus spretus*. Following *Dra I* digestion, the PRI probe detected a 6kb band in C57BL/6 DNA and a 5.5kb band in *Mus spretus* DNA (figure 3.3). Segregation of these alleles was tested in 28 backcrossed animals. The results are presented in table 3.2. Of the previously analysed markers the segregation pattern of the *Esterase 3* (*Es-3*) locus was most similar to that of pL7, with only one difference in 28 individuals tested (ie. 1/28 recombinations between the two markers). *Es-3* is at the distal end of chromosome 11 (Green, 1981). Although the number of animals examined was too small to

estimate an accurate map distance between pL7 and *Es-3*, these data indicate that pL7 is linked to *Es-3* and therefore it lies on chromosome 11. From the chromosomal location of pL7 and its sequence similarity to *Hox 1.6* (see below), it was suggested that pL7 is a new member of the *Hox 2* cluster; *Hox 2.9* (Murphy *et al.*, 1989). This has subsequently been clearly demonstrated in the human and the mouse (Acampora *et al.*, 1989; Rubock *et al.*, 1990)).

Table 3.2. The inheritance of the *Mus spretus* haplotype in backcrossed animals ([C57BL/6 X *Mus spretus*]F_1 X C57BL/6).

DNA	lambda 7 probe	Es-3 probe	DNA	lambda 7 probe.	Es-3 probe
151	+	+	172	-	-
152	+	+	173	-	-
153	+	+	178	-	-
155	+	+	180	+	+
156	-	-	187	-	-
157	+	+	188	+	+
159	-	-	189	+	+
160	-	-	190	+	+
161	-	-	192	-	-
162	+	-	196	+	+
163	+	+	197	-	-
164	+	+	198	+	+
165	-	-	199	+	+
166	+	+	200	+	+

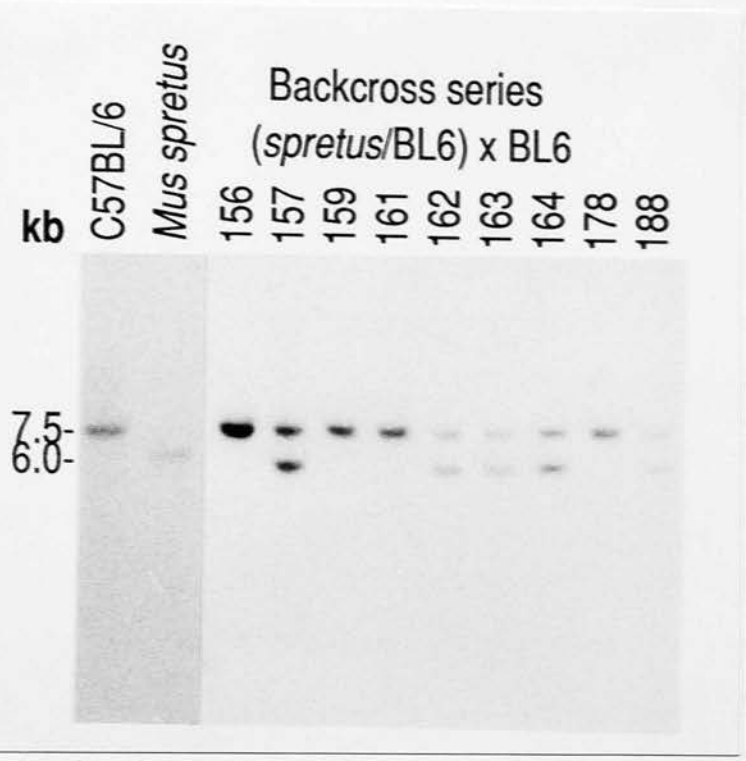


Figure 3.3. The segregation of a *Dra* I polymorphism detected by a *Hox 2.9* probe in C57BL/6 X *Mus spretus* backcrossed animals. Lanes 1 and 2 show the bands detected in *Dra* I cut parental DNA by a *Hox 2.9* specific probe. The other 9 lanes contain *Dra* I-cut DNA from backcrossed hybrid offspring. Animals 157, 162, 163, 164 and 188 inherited both *Mus spretus* and C57BL/6 chromosomes whereas the remaining animals inherited two C57BL/6 chromosomes.

3.2.3 Sequence of the Hox 2.9 *cDNA.*

The *Hox 2.9* cDNA was sequenced using three strategies in parallel (figure 3.4). Firstly, double stranded sequencing was used to sequence from both ends of specific subfragments cloned into the plasmid vector Bluescribe. The universal and reverse sequencing primer sites in the vector were used. Secondly, the 5' *Eco* RI / *Hinc* II fragment was further digested with restriction enzymes (*Alu* I, *Dde* I, *Hae* III) that recognise 4bp sites and therefore cut frequently. These fragments were randomly cloned into M13, where they were prepared as single strand molecules and sequenced from the M13 sequencing primer site. Thirdly, specific 16-20bp oligonucleotides, complimentary to areas in which the sequence was already known, were synthesised and were used to prime sequencing reactions across gaps in the known sequence or across areas where there were discrepencies in the data already obtained. In this way all areas were sequenced independently in

both orientations at least twice (figure 3.4). The individual sequences were aligned using the Amersham Staden Plus suite of computer programs.

The sequence of the 1780 bp *Hox 2.9* cDNA was found to contain a single long open reading frame of 894bp with an in-frame homeobox domain (figure 3.5). It is predicted to encode a 32kD protein. A *Hox 2.9* transcript of approximately 2kb was detected in RNA prepared from 9 day embryos and F9 teratacarcinoma cells (section 3.2.5), indicating that the sequenced cDNA clone represents an almost full length transcript. The absence of a poly(A) stretch and a recognisable polyadenylation site (Proudfoot & Brownlee, 1976) indicates that the cDNA insert is truncated at the 3' end. The predicted ATG translational start codon is the first ATG in-frame with the homeobox. No other in-frame ATG was found within 514bp upstream of the proposed start site.

Sequence comparison reveals that *Hox 2.9* contains a homeodomain similar to that of the *Drosophila labial* gene (figure 3.6A). The *Drosophila labial* homeodomain has diverged significantly from the *Antennapedia* sequence (67% amino acid identity, table 3.3) and is most closely related to its homologues in other species including *Hox 2.9* (Fig. 3.6; table 3.3). The *Hox 2.9* homeodomain shows 80% amino acid identity to the *Drosophila labial* homeodomain and 87% identity to that of *Hox 1.6*, which has previously been shown to be *labial*-like (85% amino acid identity) (Mlodzik *et al.*, 1988; LaRosa & Gudas, 1988b). Comparison of the *labial*-like homeodomains to other genes in the mouse shows at best 62% identity. It is therefore suggested that *Hox 2.9* is the second member of the *labial* subfamily in the mouse. No other mouse homeobox-containing gene has been reported which belongs to this subfamily, however, genes isolated from chicken (*Ghox lab*) (Sundin *et al.*, 1990) and human (*HOX 2I*) (Acompara *et al.*, 1989) are *labial*-like.

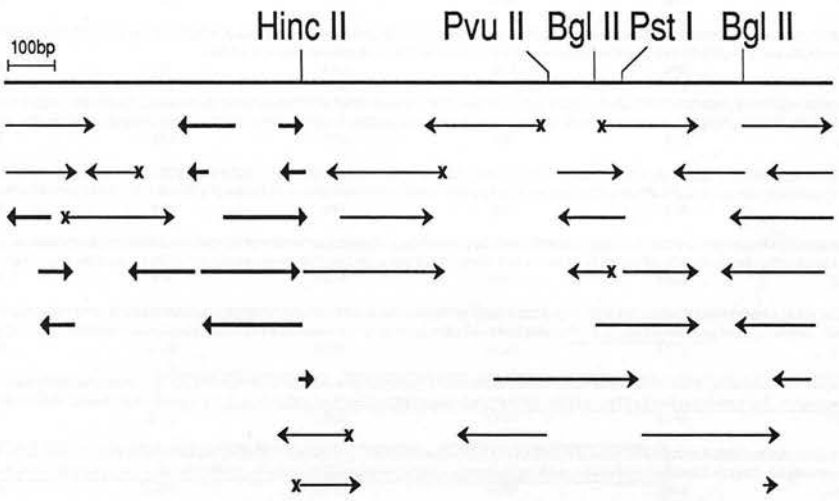


Figure 3.4 Diagramatic representation of the strategy used to sequence a *Hox 2.9* cDNA clone. Arrows represent the individual sequences that were obtained and the direction in which they were read. Anchored arrows represent regions that were sequenced from specific internal oligonucleotide primers in double-strand clones. Heavy arrows represent random restriction fragments, sequenced as single strand M13 clones. Light arrows represent sequences obtained by sequencing from the ends of specific, restriction fragment generated, double stranded, subclones.

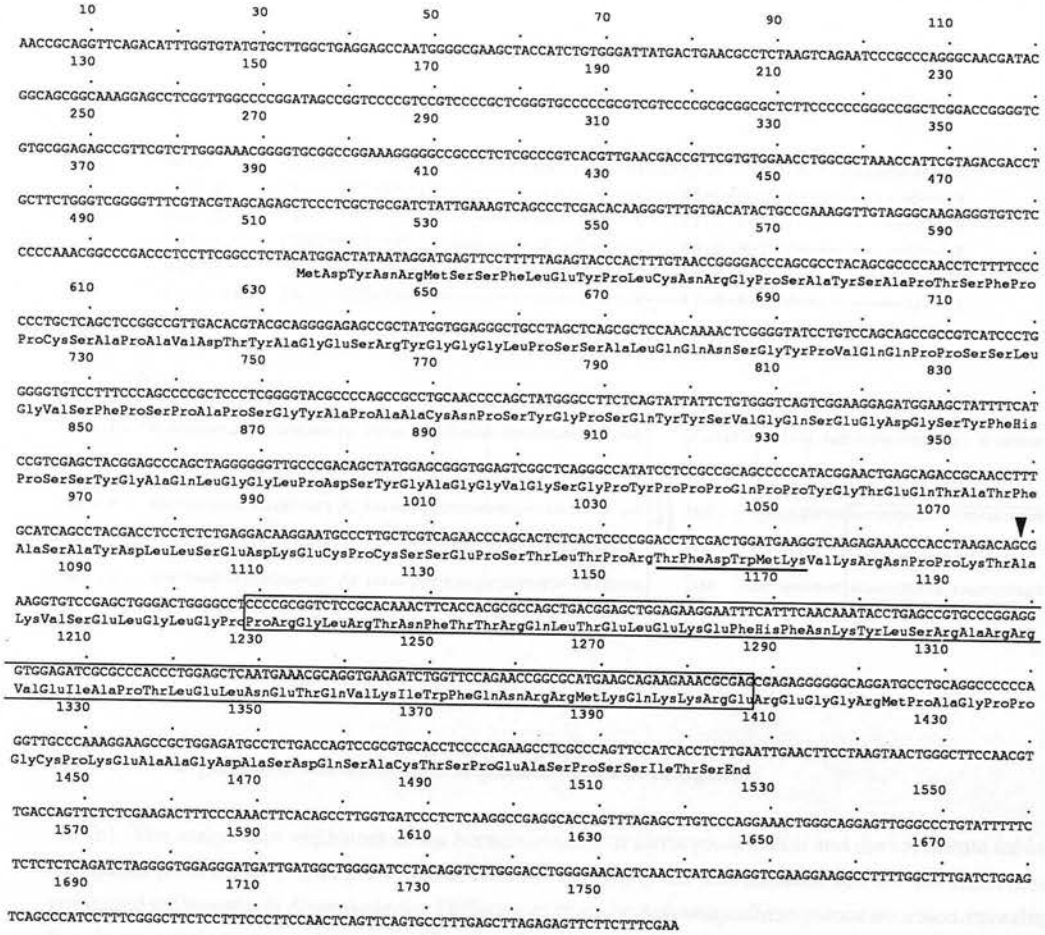


Figure 3.5 The sequence of a 1759bp *Hox 2.9* cDNA clone. The amino acid sequence of the longest open reading frame with the homeodomain in frame (boxed) is given below the cDNA sequence. The conserved hexapeptide is underlined. The arrow head shows the position of the predicted splice site based on comparison with other *labial*-like gene sequences.

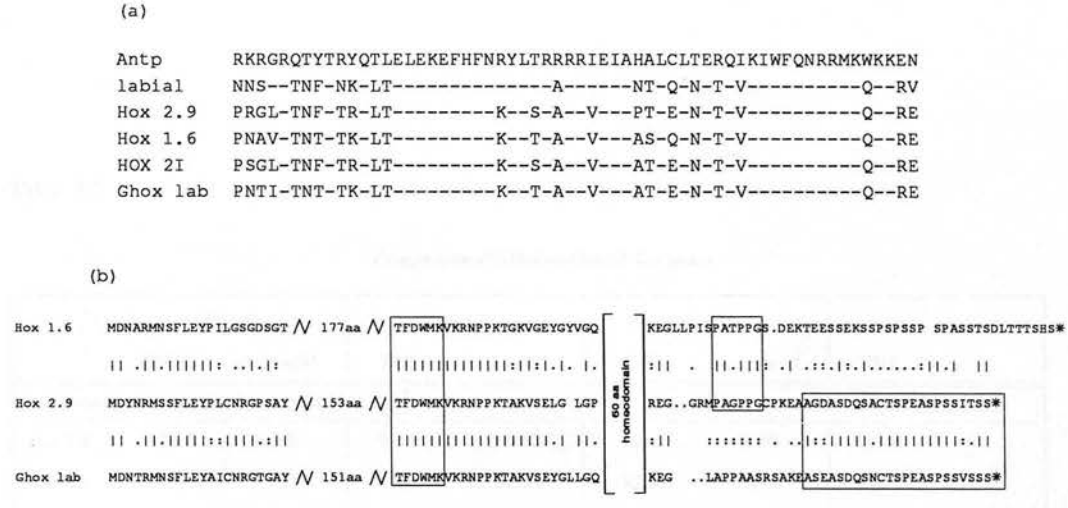


Figure 3.6 Alignment of the amino acid sequences of *labial*-like genes.

(a) The amino acid sequences of the homeodomains of *Drosophila labial* and the vertebrate *labial*-like genes (*Hox 2.9* and *Hox 1.6* : mouse, *Hox 21*: human, *Ghox lab*: chicken, see text for references) compared to *Drosophila Antennapedia*. Differences from the *Antennapedia* sequence are noted, revealing the characteristic features of the *labial* family of genes.

(b) A representation of the full coding regions of *Hox 1.6, Hox 2.9* and *Ghox lab*. A line between two sequences represents an amino acid identity. Dots represent conservative changes (as judged by the GAP program). The conserved hexapeptide is boxed. A divergent stretch 5' of the hexapeptide, and the homeodomain shown in figure 3.6a, are ommitted from this diagram.

Table 3.3

Comparison of* labial *and* labial-*like genes

	HOX 2I		*Hox 1.6*		*Ghox lab*		
	DNA	amino acid	DNA	amino acid	DNA	amino acid	DNA
(a) Homeobox / homeodomain comparisons.							
Hox 2.9	87%	96.6%	79.0%	86.6%	80.6%	90.0%	72.3%
HOX 2I			82.8%	88.3%	82.8%	91.7%	74.5%
Hox 1.6					81.1%	93.3%	76.2%
Ghox lab							74.5%
(b) Whole protein comparisons.							
	similarity	Identity	similarity	Identity	similarity	Identity	
Hox 2.9	90.3%	85.6%	62.5%	45%	71.1%	55.1%	
Hox 1.6					58.8%	46.5%	

Comparison of the mouse and chicken *labial*-like genes reveals that regions of similarity also exist outside the homeodomain (figure 3.6B). These include a stretch of 22 amino acids at the N-terminus of the proteins and regions that extend from both ends of the homeodomain. The *Hox 1.6* gene has previously been shown to contain only two amino acids (Trp-Met) of the conserved hexapeptide (Ile/Val-Tyr-Pro-Trp-Met-Arg) found in many homeodomain proteins (Baron *et al.*, 1987). However the four replacements (Trp-Phe-Asp-X-X-Lys) in this region in *Hox 1.6* are conserved in both *Hox 2.9* and *Ghox lab* and interestingly, mark the beginning of the extended region of homology around the homeodomain. Examination of the full coding region shows that *Ghox lab* is more similar to *Hox 2.9* (figure 3.6B; table 3.3), particularly at the C-terminal end. *Ghox lab* and *Hox 2.9* also have the same pattern of expression in the developing embryo (Sundin & Eichele, 1990; Murphy & Hill, 1991) indicating that these two genes are homologues.

3.2.4 Restriction mapping of a Hox 2.9 *genomic clone.*

In order to analyse the complete *Hox 2.9* gene, a *Hox 2.9* clone was isolated from a mouse genomic library in lambda 2001 (library received from T. Rabbits). The clone (L2.9) is approximately 18.5kb long and contains the full coding region of *Hox 2.9*. Restriction sites within the clone were mapped in two ways:

(1). Partial digests of L2.9 were prepared with four enzymes (*Xba* I, *Sac* I, *Xho* I and *Eco* RI) that do not cut in the vector; *Xba* I, *Sac* I and *Xho* I cut in the polylinker region into which the insert was cloned, while the *Eco* RI polylinker site was removed by the cloning procedure. A time course of digestions was carried out to ensure that all possible partial digests were included. The digests were separated on low percentage (0.6%) agarose gels in order to resolve large fragments. The gels were Southern blotted and probed with an oligonucleotide complimentary to the lambda short arm. This detected a variety of fragments, the fragment length depending on the position of the particular site within the clone that was cleaved (figure 3.7). The position of the restriction sites could therefore be worked out from the length of the fragments produced.

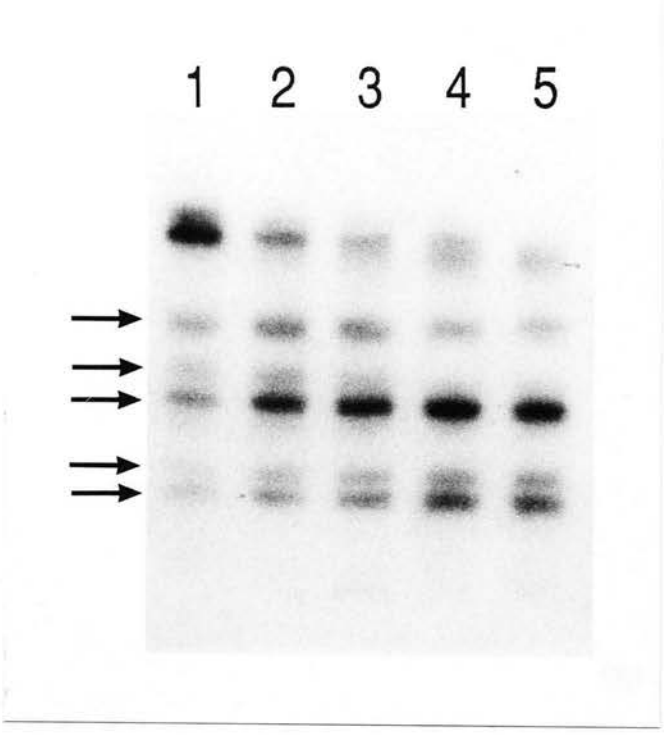


Figure 3.7. lambda 2.9, a *Hox 2.9* genomic clone partially digested with *Eco* RI, separated on a 0.6% agarose gel, Southern blotted and probed with an oligonucleotide specific to the lambda short arm. The five partial bands indicated by the arrows on the left represent the five internal *Eco* RI sites indicated in figure 3.8. Each band was created when the appropriate site within the clone was cleaved and the position of the site was worked out from the length of the fragment. *Xba* I, *Sac* I and *Xho* I sites were similarily mapped.

(2). Two anchor restriction sites were used to map sites in and around the *Hox 2.9* gene. These were *Eag* I, which falls within exon 1, and *Xho* I, which is located 1.2kb 3' of exon 2. Information was obtained on the positions of restriction sites for 13 different enzymes relative to these two anchor sites. This was achieved by screening complete single and double (plus the anchor site enzyme) digests with probes from the *Hox 2.9* coding region, both 3' and 5' of the intron.

The restriction mapping data are compiled in figure 3.8. The relative distances of sites in exon 1 from sites in exon 2 reveal that the *Hox 2.9* coding sequence is interrupted by an intron of approximately 400bp. The position of the intron was predicted from comparison with other *labial*-like genes and this was verified by the work of Frohman *et al* (1990).

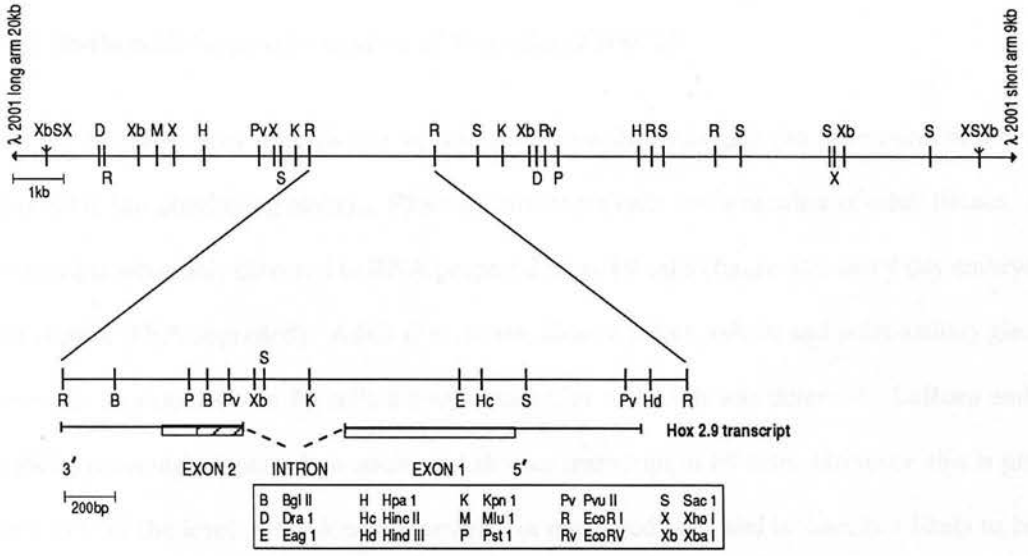


Figure 3.8 A restriction site map of an 18.5kb *Hox 2.9* genomic clone.

The L2.9 genomic clone contains more than 10kb of DNA upstream of the *Hox 2.9* coding sequence, the entire *Hox 2.9* transcription unit (including the intron) and approximately 6kb of 3' flanking DNA. It is likely therefore that the clone contains most, if not all, of the *Hox 2.9* controlling sequences (see section 1.4.3). This clone can be used in the future to produce gene fusions with reporter genes to assay *Hox 2.9* reporter activity (see section 6.2).

3.2.5 Northern blot expression analysis of Hox 1.6 *and* Hox 2.9.

Northern blot analysis was used to preliminarily investigate the expression of *Hox 1.6* and *Hox 2.9* in the developing embryo, F9 teratocarcinoma cells and a number of adult tissues. *Hox 1.6* transcripts were only detected in RNA prepared from F9 cells (figure 3.9) and 9 day embryos (data not shown, RNA degraded). Adult liver, brain, kidney, testes, spleen and submaxillary gland were found to be negative. In F9 cells a single transcript of ~2.1kb was detected. LaRosa and Gudas (1988b) previously reported an additional shorter transcript in F9 cells. However, this is present at only 10% of the level of the longer transcript in untreated cells and is therefore likely to be below the level of detection in this experiment. There is an approximately 20-fold induction in the level of the *Hox 1.6* transcript following a 24 hour exposure of F9 cells to 5×10^{-7} M RA. This is seen in both poly(A) and total RNA preparations; in total RNA the transcript only becomes detectable after RA treatment.

A *Hox 2.9* transcript was detected in both F9 cells treated with RA and in 9 day embryos (figure 3.10). The single transcript band runs just below the level of 18S ribosomal RNA and is therefore approximately 2kb long. An induction of *Hox 2.9* expression following treatment with RA, similar to that seen with *Hox 1.6*, is suggested by the presence of a band only in F9 cell RNA after treatment with RA. It also indicates that the level of expression of *Hox 2.9* in F9 cells is lower than *Hox 1.6*, both before and after treatment. A clear band is detectable in 9 day embryonic RNA but not in RNA from older embryos, indicating that the gene is transcribed only in the early embryo.

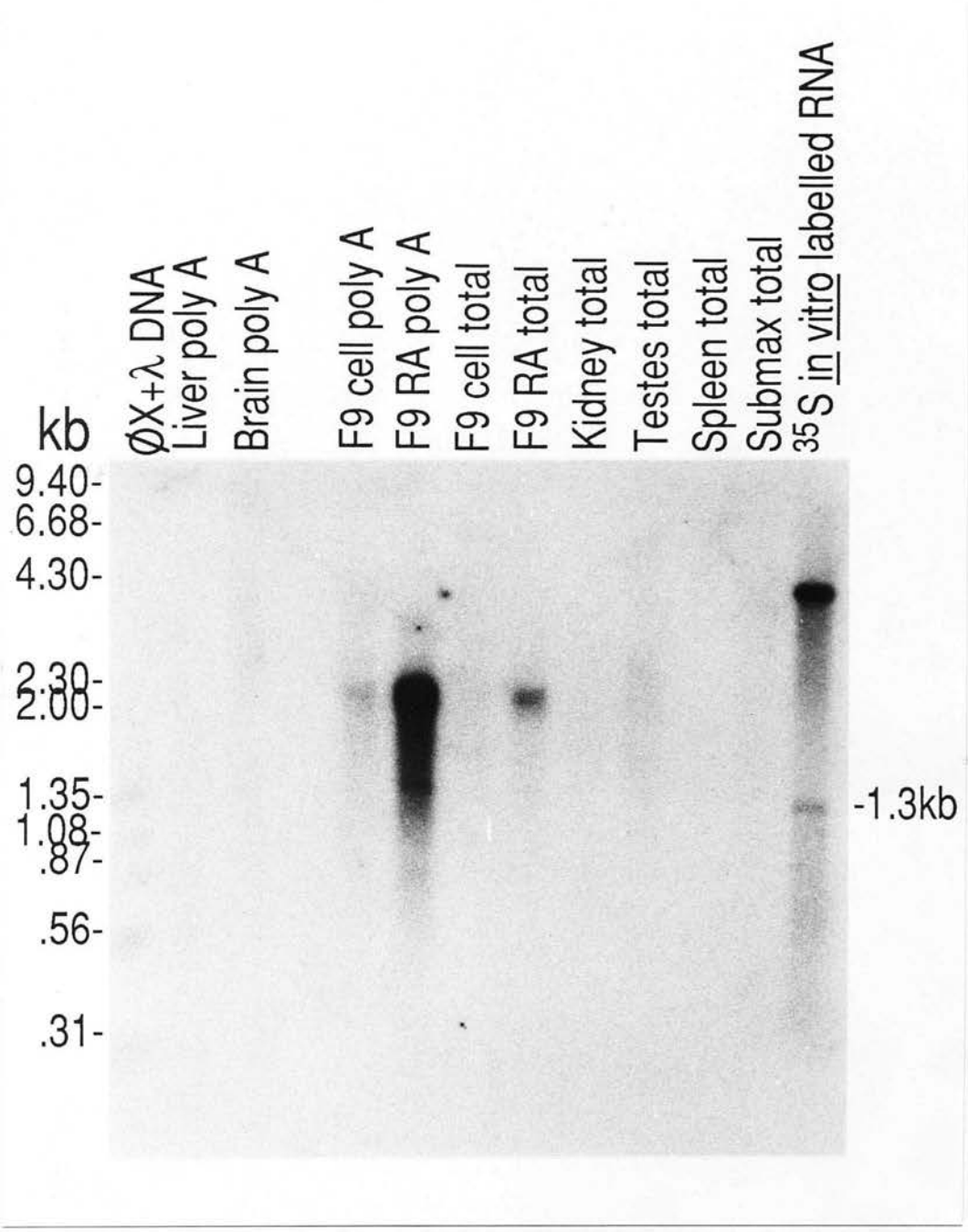


Figure 3.9. A Northern blot probed with a *Hox 1.6* specific probe showing expression in F9 teratocarcinoma cells. An approximately 20-fold induction was observed following exposure of the cells to 5×10^{-7} M RA for 24hrs.

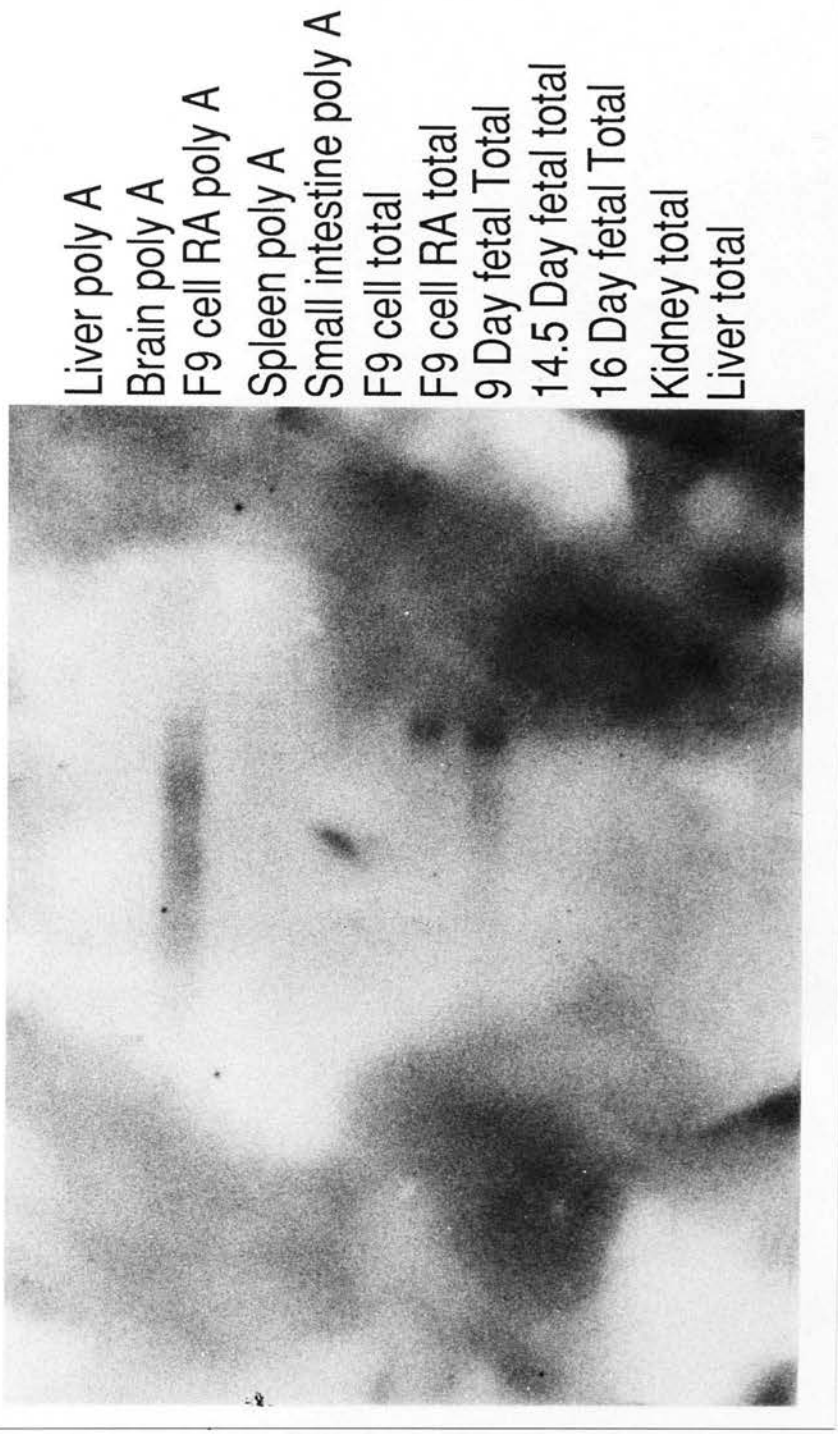


Figure 3.10. Northern blot analysis of Hox 2.9 expression showing a 2kb transcript in total RNA from F9 cells that were exposed to RA. A RA induction is indicated by the fact that no transcript was detected in untreated cells. A transcript was also detected in RNA from 9 day embryos but not at later stages.

Hox 2.9 expression in the early embryo was further investigated by *in situ* hybridisation. This is reported in the following section. Further analysis of the expression of *Hox 1.6* is reported in chapter 4.

3.2.6 In situ *hybridisation analysis of* Hox 2.9 *expression.*

The pattern of *Hox 2.9* expression during development was investigated by *in situ* hybridisation. This technique involves the hybridisation of whole embryo sections with synthetically produced, radioactively labelled, antisense RNA that can recognise the endogenous transcripts 'in situ'. Therefore it reveals the genes temporal and spatial pattern of expression. The PRI subclone of *Hox 2.9* was used to make ^{35}S labelled sense and antisense RNA by initiating transcription from the T3 and T7 promoters in the Bluescribe vector. In this initial study, serial sections from $6^1/_2$ day, $8^1/_2$ day, $9^1/_2$ day, 10 day and $13^1/_2$ day embryos were analysed.

Hox 2.9 transcripts were not detected in the $6^1/_2$ day old embryo, at the time that gastrulation begins. At $8^1/_2$ days of development the embryo is at the neural fold stage and expression of *Hox 2.9* is detected in both the neural-ectoderm and the underlying mesoderm (figure 3.11). The domain is extensive, with an anterior limit in the developing hindbrain. It proved difficult to prepare a section that was continuously sagittal along the length of the embryo at this stage. This appears to be due to the process of embryonic rotation which has begun to take place (see figure 1.1). Serial sections indicated that the expression domain is continuous along the trunk of the embryo but this is not obvious in the section shown in figure 3.11. This was established more clearly in the set of experiments described in chapter 4. The low level of expression and the absence of morphological markers in the hindbrain in this experiment made it difficult to precisely locate the anterior limit of expression. Mesodermal expression was restricted to lateral and presomitic mesoderm; no label was detectable in the somites.

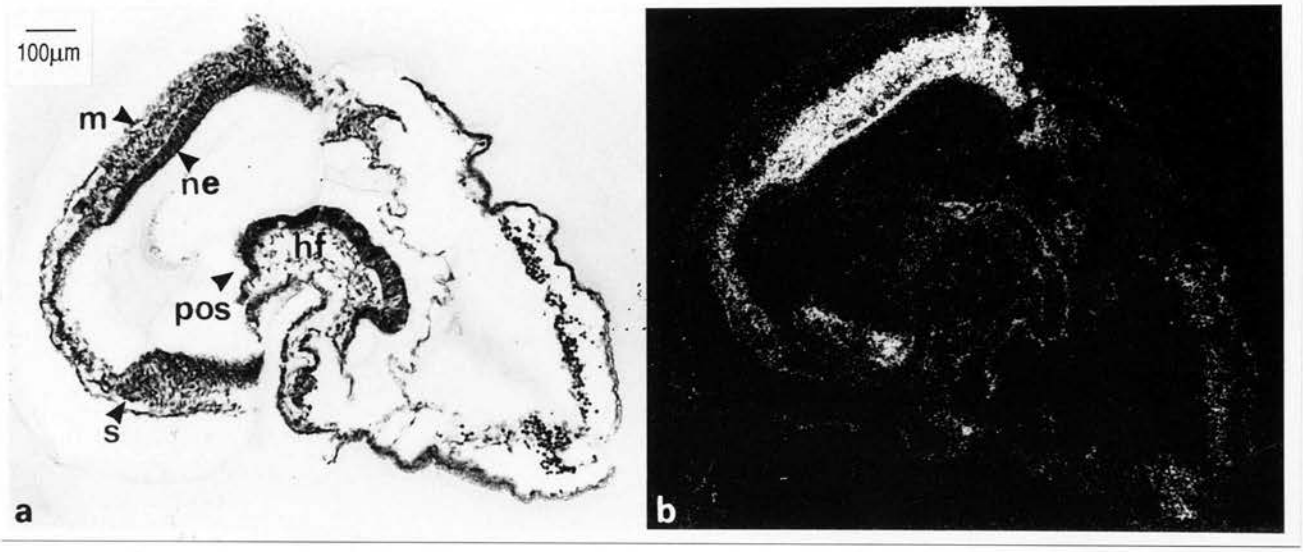


Figure 3.11. *In situ* hybridisation of ^{35}S labelled antisense probe for *Hox 2.9* to an $8^1/_2$ day mouse embryo section. a, shows a bright-field image and b, a dark-field image. m, mesoderm; ne, neuroectoderm; pos, preotic sulcus; hf, headfold.

By $9^1/_2$ days the expression pattern had changed dramatically, as had the complexity of the embryo (figure 1.1). At this stage the segmental units of the hindbrain, the rhombomeres, could easily be seen as a series of neuroepithelial swellings in whole embryos (figure 1.2) and in frontal and parasagittal sections (figure 3.12A and C). Individual rhombomeres could be identified by their position relative to the otic vesicle and are numbered as shown in figure 3.12. In the hindbrain, Hox 2.9 was expressed at this stage in a single segment, rhombomere 4. It was difficult, however, to relate rhombomere 4 to the hindbrain expression at day $8^1/_2$, when rhombomeres are not morphologically detectable. This problem is addressed in chapter 4. The *Hox 2.9* boundaries of expression correspond precisely with the morphological boundaries of rhombomere 4 (figure 3.12A and C). Figure 3.12D shows the precision of this boundary at the cellular level where all the cells within the segment appear labelled, and the labelling ceases abruptly in a straight line at the

segment boundary. Rhombomere 4 seems to be the principle site of expression of *Hox 2.9*, because the very localised signal detected at this site was consistently higher than any other signal detected.

Outside the hindbrain the anterior neural tube was unlabelled at $9^1/_2$ days, but expression was detected in the neural tube posterior to the forelimb buds (figure 3.13). Expression was still detected in the presomitic mesoderm but not in the somites (figure 3.12E and F). Gut-associated mesenchyme and the epithelium of the upper gut posterior to the third branchial arch were also labelled (figure 3.13), the gut-associated mesoderm being derived from lateral plate mesoderm, which showed expression at $8^1/_2$ days. A low level of expression was detected in the nephrogenic duct (figure 3.13). Analysis of sections through 10 day embryos showed no further change in *Hox 2.9* expression (figure 3.13). Unlike most homeobox-containing genes previously analysed, the expression of *Hox 2.9* did not persist at later embryonic stages: at $13^1/_2$ days of development, expression could not be detected in any part of the embryo by examination of serial sections in sagittal, transverse and frontal planes. Neuromeric boundaries can also no longer be seen at this stage.

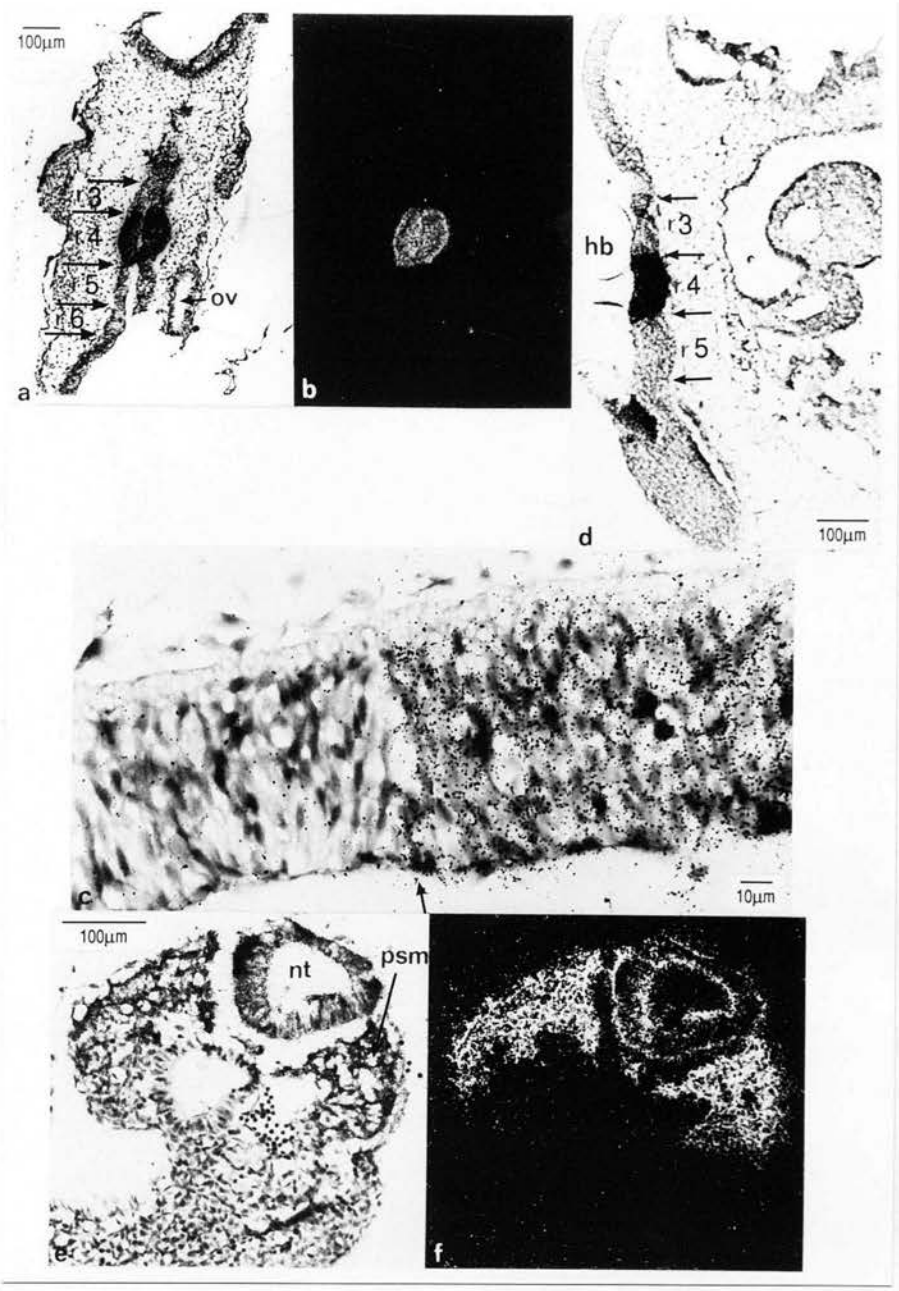


Figure 3.12. Expression of *Hox 2.9* in the $9^1/_2$ day embryo. a-b, show a frontal section through the hindbrain. c, is a high magnification photomicrograph of the boundary between rhombomeres 3 and 4 (marked with arrow). The short autoradiographic exposure time (two weeks) allowed the labelling of individual cells to be observed and emphasised the sharpness of the expression boundary. d, shows a sagittal section probed with *Hox 2.9* and exposed for a long period (5 weeks). e-f show a transverse section through the posterior region of the embryonic trunk. b and f are darkfield images of a and e respectively. r3-r6, rhombomeres 3 to 6; ov, otic vesicle; hb, hindbrain; nt, neural tube; psm, presomitic mesoderm.

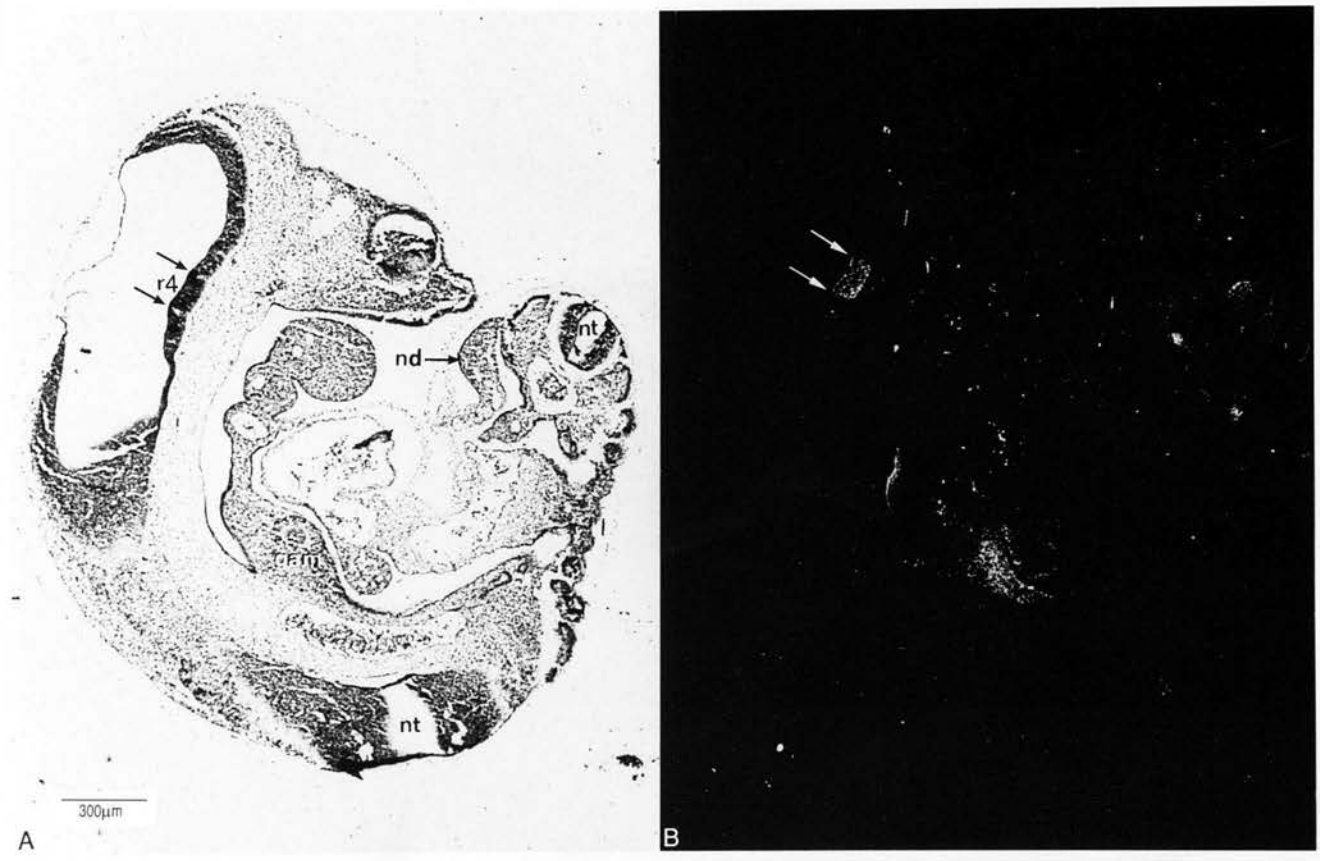


Figure 3.13. Expression of *Hox 2.9* in the 10 day embryo. b is a darkfield image of the sagittal section shown in a. r4, rhombomere 4; nt, neural tube; nd, nephrogenic duct; gam, gut associated mesoderm.

3.3 Discussion.

3.3.1 The Hox 1.6 *transcript is differentially spliced.*

In F9 teratocarcinoma cells two alternate *Hox 1.6* transcripts are produced by differential splicing (LaRosa & Gudas, 1988b). In uninduced F9 cells, the larger transcript is about 10 times more abundant than the shorter transcript. This relative abundance varies with time of exposure to RA. Here, the relative proportions of the two transcripts in an $8^1/_2$ day embryonic cDNA library were investigated and it was found that the shorter transcript was represented in 7/17 clones (~40%).

The alternate transcripts differ by the presence or absence of a 203bp sequence 5' of the homeobox (figure 3.1). The sequences surrounding this region show some similarity to consensus splice site sequences (Mount, 1982). The presence of imperfect splice sites could explain why this region is spliced out in only some transcripts and suggests that regulation of the proportion of alternate transcripts may occur by modulation of the capacity of the splicing machinery to recognise these sites. The existence of such a regulatory mechanism is suggested by the different proportions of transcripts found in the $8^1/_2$ day embryo and in F9 cells before and after RA treatment. This feature of *Hox 1.6* was subsequently further investigated and is reported in chapter 4.

The alternatively spliced region of *Hox 1.6* falls within the portion of the two mouse *labial*-like genes that are most divergent (figure 3.6B). There is no evidence, either from sequence analysis or from detection of transcripts, that a similar differential splicing mechanism operates on the *Hox 2.9* gene. Frohman *et al.* (1990) have further investigated this possibility by PCR amplification across the equivalent region of *Hox 2.9.* They found that only a single sized band was amplified, which suggests that *Hox 2.9* is not differentially spliced, at least in this region.

The detection of different transcripts of homeobox-containing genes is not uncommon. For example, in the *Drosophila* system the *Antennapedia* gene has multiple transcriptional start sites driven by different promoters (Carroll *et al.*, 1986; Schneuwly *et al.*, 1986). The *Ultrabithorax* gene produces many different transcripts by differential splicing of several small exons (Sanchez-Herrero *et al.*, 1985; Ingham, 1985; Akam & Martinez-Arias, 1985). In *Xenopus*, there are two alternative *XlHbox1* transcripts produced by different promoters (Cho *et al.*, 1988). The *XlHbox1* encoded proteins are identical with the exception of an 82 amino acid extension at the amino terminal of one transcript. Different transcriptional start sites have also been reported for the mouse gene *Hox 2.1* (Krumlauf *et al.*, 1987). However, the situation described for *Hox 1.6*, where the removal of upstream sequences leads to a shift in the reading frame and the production of an alternative transcript that can not encode a homeodomain, has not been observed elsewhere.

3.3.2 Evolution of the labial *family of genes.*

It was shown here that *Hox 2.9* together with *Hox 1.6* (Mlodzik *et al.*, 1988; LaRosa & Gudas, 1988b) are the mouse homologs of the *Drosophila labial* gene. *labial*-like genes have also been identified in the human and chicken. The homeodomains of the vertebrate genes are very similar to that of *labial* (80-85% identical, table 3.3), but there is little similarity throughout the rest of the protein. The *Drosophila* protein is 629 amino acids long whereas the vertebrate proteins are much shorter, varying between 298 and 336 amino acids. Vertebrates also lack the intron that interrupts the homeodomain of the *Drosophila* gene (Diedrich *et al.*, 1989; LaRosa & Gudas, 1988b; Acampora *et al.*, 1989). The vertebrate genes are more similar to each other in structure and sequence, with homology extending outside the homeodomain (figure 3.6B). These regions of extended homology indicate the regions of the proteins that are important for common *labial*-like gene functions. Comparing the full length proteins it appears that the chicken gene *Ghox-lab* is more similar to *Hox 2.9* than *Hox 1.6* although the homeodomains of all three genes are very similar. Expression analysis also indicates that *Ghox-lab* is the homologue of *Hox 2.9*.

The segmental expression of *Ghox lab* in the hindbrain (Sundin & Eichele, 1990) is valuable since the chicken is a useful system for developmental manipulation.

The genetics of the *Drosophila* gene *labial* have proven difficult to interpret, but a homeotic role for *labial* has been concluded from clonal studies (Merrill *et al.*, 1989). The *labial* protein has been found in neural and epidermal cells of a very distinct region of the head which is thought to represent an ancestral segment (Diedrich *et al.*, 1989). The fact that *labial* and one of its mouse homologs, *Hox 2.9,* are expressed in single anterior segments is striking. Although it is likely that there are differences in the systems for determining position in two such distinct and specialised organisms, these highly conserved genes are involved in both.

As a result of duplication there are at least two *labial*-like genes in the mouse, and most likely in other vertebrates. The duplicated genes have diverged but have maintained certain features to a remarkably high degree, which in itself would indicate a conservation of function at some level. The two mouse genes also share several features of their expression patterns (chapter 4) which extends the theory of functional similarity. The duplication of these genes must have been necessary to accommodate the establishment of the more complex body plan of the vertebrate in which members of subfamilies maintain similar although specialised functions.

3.3.3 Hox 2.9 *and* Hox 1.6 *are expressed early in development and are induced following RA treatment of F9 cells.*

Northern blot analysis reveals that *Hox 1.6* and *Hox 2.9* are both expressed at 9 days of development. Furthermore, *Hox 2.9* is not detected in the later embryo ($14^1/_2$ days). It was previously shown by Northern blot analysis that *Hox 1.6* is also only detected at early stages (Baron *et al.*, 1987). This is in contrast to all other *Hox* genes analysed to date (Holland & Hogan, 1988b; Graham *et al.*, 1989), which reveals a notable similarity in the expression of the two *labial*-like genes. The expression patterns of these two genes are compared in detail in chapter 4.

When F9 teratocarcinoma cells in monolayer culture are exposed to RA they are induced to differentiate toward parietal endoderm (Strickland & Mahdavi, 1978). This was demonstrated by the induced expression of characteristic genes (Strickland *et al.*, 1980; Wang *et al.*, 1983; Wang *et al.*, 1985). The expression of homeobox genes has also been shown to be altered in RA-treated cells (Hauser *et al.*, 1985; Colberg-Poley *at el.*, 1985; LaRosa & Gudas, 1988b; Simeone et al., 1990). The Northern analysis presented here shows that the level of both *Hox 1.6* and *Hox 2.9* RNA in F9 cells is increased following RA treatment. Such an increase in the level of transcripts may be brought about in a number of ways. Firstly, there may be stabilisation of the transcript. However, LaRosa and Gudas (1988b) have shown that this is not the case for *Hox 1.6.* Secondly, there may be indirect transcriptional induction, as is the case for some of the endodermal marker genes (Gudas & Wang, 1986). In this case induction is dependent on protein synthesis and is blocked by cycloheximide. This again was shown to be untrue for *Hox 1.6* (LaRosa & Gudas, 1988a), and other homeobox genes are induced rapidly indicating that the induction does not require protein synthesis. Thirdly, and most interestingly, there may be primary induction through the RA receptor system (section 1.2.6). Although this has not been directly demonstrated, the characteristics of induction of at least some homeobox genes are consistent with a primary response. In chapter 5 the effect of excess RA on the expression of *Hox 2.9* in the developing embryo is investigated.

3.3.4 Hox 2.9 *is expressed in a single segmental unit of the developing hindbrain.*

In situ hybridisation analysis showed that *Hox 2.9* transcripts are spatially restricted within the early developing embryo. This is the case for all homeobox-containing genes so far analysed and is consistent with a role for these genes in positional determination (section 1.4.3). *Hox 2.9* is expressed in mesodermal, endodermal and ectodermal tissues but the most striking feature of the expression pattern is expression within a single segment of the developing hindbrain; rhombomere

4. The expression of *Hox 2.9* in rhombomere 4 indicates that the gene plays a part in the formation or differentiation of this segment.

Rhombomeres are present throughout the vertebrates indicating the importance of these transient metameric structures for the development of the hindbrain (section 1.1.4). Their organisational role is reflected in the pattern of nerve generation observed in the chick (Lumsden & Keynes, 1989; section 1.1.4). Perfect correlation between the expression of *Hox 2.9* and the morphological limits of rhombomere 4 (figure 3.) indicates that *Hox 2.9* may determine the identity of this segment, acting therefore in an equivalent manner to the structurally similar *Drosophila* homeotic selector genes (section 1.3). In the chick, where the timing and pattern of hindbrain neurogenesis has been well described (Kuratani *et al.*, 1988; Covell & Noden, 1989; Lumsden & Keynes, 1989), the chicken equivalent of *Hox 2.9* (*Ghox lab*) is expressed prior to and during the differentiation of motor neurons (Sundin & Eichele, 1990). Thus *Hox 2.9* may influence this process in the mouse.

Independent evidence for the genetic control of development within rhombomere 4 comes from the phenotype of the mouse developmental mutant *kreisler*, which displays faulty segmentation of the hindbrain and specific degeneration of the cells of rhombomere 4 by $9^1/_2$ days (Deol, 1964). The mutation maps to mouse chromosome 2 (Hertwig, 1942) and is therefore not allelic with *Hox 2.9*. It is possible that *Hox 2.9* and *kreisler* are both genetic regulators of segmentation in the hindbrain. For this reason it is of interest to investigate the relationship between these two genes. However, an attempt to analyse the expression of *Hox 2.9* in homozygous *kreisler* embryos was hampered by difficulties in breeding *kreisler* mutants in this laboratory.

Krox 20 is expressed in rhombomeres 3 and 5 (Wilkinson *et al.*, 1989a) which flank the expression domain of *Hox 2.9* at 9 days. *Krox 20* is also thought to be a regulatory gene on the basis of a zinc finger DNA binding motif contained within the encoded protein (Chavrier *et al.*, 1988). Proteins with zinc-finger and helix-turn-helix DNA-binding domains interact during

Drosophila segmentation (Ingham *et al.*, 1988). The spatial and temporal correlation between the expression of *Hox 2.9* and *Krox 20* raises the possibility that their products interact directly or indirectly in the process of segmentation of the mouse hindbrain. The expression domains do not appear to overlap however, indicating that interaction is only possible at segment boundaries. Alternatively, expression may overlap transiently as the domains become established. The more detailed temporal analysis of both genes presented in chapter 4 failed to detect simultaneous expression of the two genes, but shows that cells within rhombomere 5 express *Hox 2.9* prior to expressing *Krox 20.*

Hox 2.9 is the only known homeobox-containing gene to be expressed in a single hindbrain segment. However, the other members of the *Hox 2* cluster, which occupy more extensive overlapping domains, have discrete anterior boundaries within the CNS, with those of the more anteriorly expressed genes corresponding with rhombomere boundaries (Wilkinson *et al.*, 1989b). It was previously suggested that the overlapping domains of homeobox-containing gene expression convey positional information in the form of unique combinations of gene products in blocks along the AP axis (Holland & Hogan, 1988b). It appears that within the hindbrain these blocks correspond to two adjacent segments, with progressively more 3' genes occupying boundaries two segments more anterior than the previous gene. *Hox 2.9* does not follow this pattern. This is yet another feature of *Hox 2.9* that demonstrates its highly distinctive nature, and most likely the distinctive nature of the *labial*-like subclass in general.

Chapter 4

Expression of the mouse labial-like homeobox-containing genes, Hox 2.9 and Hox 1.6, during segmentation of the hindbrain.

4.1 Introduction.

In chapter 3 the division of mouse homeobox-containing genes into subfamilies was described (section 3.1). The subfamilies are based on sequence homology and where the genes are part of gene clusters, on the relative position of the genes within the clusters. The subfamily members are therefore thought to have arisen from a common ancestoral gene following chromosomal duplication events. It was also pointed out that the mouse clusters can similarily be aligned with the insect HOM cluster. At one end of the insect cluster resides the gene referred to as *labial* (Diederich *et al.*, 1989; Mlodzik *et al.*, 1988). At corresponding positions in the mouse *Hox 1* and *2* clusters are the homologous genes *Hox 1.6* (Baron *et al.*, 1987, LaRosa & Gudas, 1988b) and *Hox 2.9* (Rubock *et al.*, 1990). The structural similarity between the genes has already been described (chapter 3). Since *labial*-like counterparts in the other two gene clusters have not been reported, *Hox 1.6* and *Hox 2.9* form the *labial*-like subfamily of homeobox-containing genes in the mouse.

It appears that subfamily members in different clusters display similar, although not always identical AP expression domains. This was observed for the domains of *Hox 3.3* (formerly *Hox 6.1*) and *Hox 1.2* (Gaunt *et al.*, 1988) and for the domains of *Hox 1.4, Hox 2.6* and *Hox 5.1* (Gaunt *et al.*, 1989) in the CNS and prevertebral column at $12^1/_2$ days. *Hox 1.5* and *Hox 2.7* both have anterior boundaries within the CNS that correspond to the same rhombomere boundary, the anterior boundary of rhombomere 5 (Gaunt *et al.*, 1987; Wilkinson *et al.*, 1989a). However, two separate studies on the related pair of genes *Hox 2.5* (Graham *et al.*, 1989) and *Hox 5.2* (Duboule & Dolle, 1989) show that expression of the former extends more anteriorly. This has been interpreted as indicating that similarity between paralogues does not hold for genes expressed only in the posterior embryo (Gaunt *et al.*, 1989). Some mouse homeobox-containing genes are not contained within clusters and these also have closely related genes on other chromosomes. Detailed studies of these genes, the *engrailed*-like mouse genes *En1* and *En2* (Joyner & Martin,

1988; Davidson *et al.*, 1988) and the *Msh*-like mouse genes *Hox 7.1* and *Hox 8.1* (Hill *et al.*, 1989; Davidson *et al.*, in preparation) show that they have overlapping or complimentary expression patterns. Chapter 3 reported similarities in the temporal expression of *Hox 2.9* and *Hox 1.6* but a more detailed expression analysis is necessary to determine the relationship between the spatial expression domains.

The preliminary expression analysis reported in chapter 3 also revealed that *Hox 2.9* is expressed in a single segmental unit (rhombomere) of the developing hindbrain at $9^1/_2$ days. This suggests that it plays an analogous role to *Drosophila* homeobox-containing genes in conveying positional information in the mouse, specifically in identifying rhombomere 4 (Murphy *et al.*, 1989). However, at an earlier stage (8-$8^1/_2$ days) a more extensive expression pattern was detected and it was not clear how these two patterns related to each other. This made it important to analyse intermediate stages to further investigate the function of *Hox 2.9* during development. We also wished to determine if the other *labial*-like gene, *Hox 1.6*, displays similar properties. Therefore, a detailed expression analysis of these two genes together with an additional rhombomere specific gene, *Krox 20* (Wilkinson *et al.*, 1989a), was carried out on consecutive embryonic sections at closely staged intervals. *Krox 20* was used as a temporal and positional molecular marker in the developing hindbrain. This analysis was specifically designed to compare the expression patterns of the *labial*-like genes in the mouse, to investigate how the segmental expression of *Hox 2.9* in rhombomere 4 becomes localised from an earlier more widespread domain and to determine how the onset of localised, segmental expression relates to the appearance of morphological segments. In addition a comparison was made of probes which distinguish the differential transcripts of *Hox 1.6*.

4.2 Results.

4.2.1 Hox 2.9 *and* Hox 1.6 *expression between* $7^1/_2$ *and 9 days of development (formation of rhombomeres)*

Between $7^1/_2$ and 9 days of development the expression patterns of *Hox 2.9* and *Hox 1.6* change rapidly and dramatically. To establish how the patterns evolve during this period we have examined embryos at closely staged intervals of approximately 6 hours, using the *Krox 20* (Wilkinson *et al.* 1989a) gene as a molecular marker for events occuring in the developing hindbrain (see materials and methods (2.7) for details of the probes and staging of the embryos, sections 2.7.5 and 2.8.1). Expression of *Hox 2.9* and *Hox 1.6* is first detected at $7^1/_2$ days during the early stages of gastrulation (figure 4.1). Both genes are expressed within the primitive streak in newly formed mesoderm and overlying ectoderm. *Hox 2.9* expression is at a higher level and is more extensive than *Hox 1.6.* In the early 8 day embryo *Hox 2.9* and *Hox 1.6* are expressed at a high level in extensive domains, extending from the posterior end of the embryo along the neuroectoderm and mesoderm tissue layers into the region of the developing hindbrain of the headfold (figure 4.2A-D). The two genes have identical, sharp anterior boundaries of expression in the neuroectoderm that coincide with the pre-otic sulcus (a characteristic groove in the surface of the presumptive hindbrain). An adjacent section probed with the *Krox-20* gene shows that it is expressed in a single domain in the hindbrain, the posterior boundary of which corresponds to the anterior boundary of *Hox 2.9* and *Hox 1.6* (figure 4.2D). As was previously described, *Krox 20* is first detected in a single domain in the hindbrain and later in its characteristic two stripe pattern (Wilkinson *et al.* 1989a). Within the mesoderm of the 8 day embryo both *Hox 2.9* and *Hox 1.6* expression is restricted to lateral plate mesoderm as far anterior as the headfold (figure 4.3D-F) and to presomitic (paraxial) mesoderm with expression decreasing as somites condense.

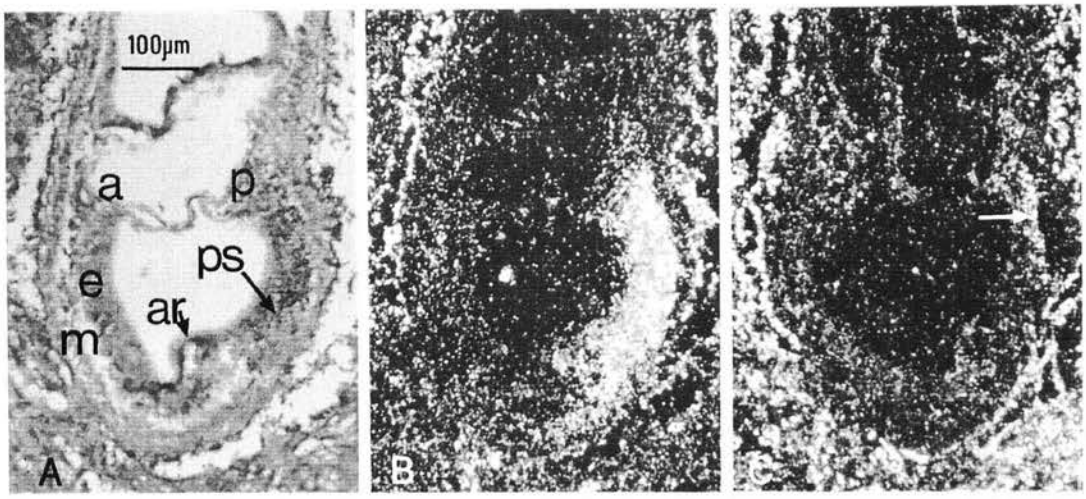
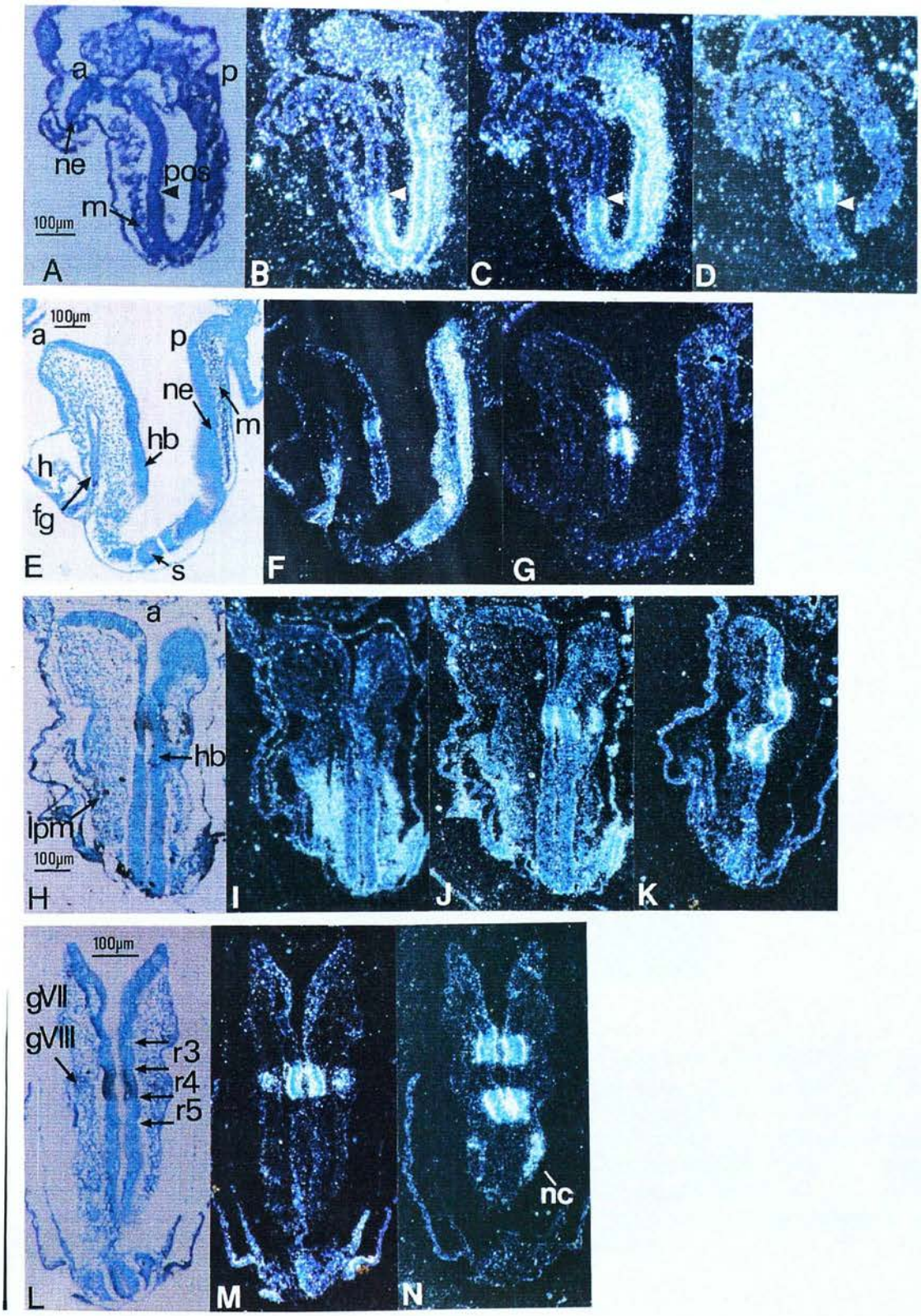


Figure 4.1. Adjacent sagittal sections through a $7^1/_2$ day mouse embryo. A and B show bright- and dark-field images of the same section probed with *Hox 2.9*. C was probed with *Hox 1.6*. a, anterior; p, posterior; ps, primitive streak; e, ectoderm; m, mesoderm; ar, archenteron. The arrows indicate the labelled cells in C and D.

By 8 1/2 days of development *Hox 2.9* expression in the hindbrain has become localised (figure 4.2E-G). The anterior boundary is at the same position as at 8 days (the pre-otic sulcus) with *Hox 2.9* and *Krox 20* continuing to share this boundary, but now there is a new posterior boundary also within the hindbrain. We simultaneously detect the initiation of the second band of *Krox 20* expression, the anterior boundary of which coincides with the posterior boundary of *Hox 2.9*. The expression of *Hox 2.9* in the anterior neural tube seems to retreat posteriorly at this time with expression persisting in more posterior regions. *Hox 1.6* expression also appears to retreat posteriorly along the neural tube in the same way, however in contrast to *Hox 2.9* no expression of *Hox 1.6* remains in the hindbrain (figure 4.2H-K). Within the mesoderm both *Hox 2.9* and *Hox 1.6* remain expressed in lateral plate mesoderm up to the level of the posterior hindbrain and in presomitic mesoderm. We now first detect expression of both genes in an endodermal derivative, the epithelium of the fore gut pocket (figure 2E+F).

Figure 4.2. Expression of *Hox 2.9, Hox 1.6* and *Krox 20* during segmentation of the hindbrain. A-D are adjacent sagittal sections from an 8 day embryo. E-G are adjacent sagittal sections from an $8^1/_2$ day embryo. H-K are adjacent frontal sections from an $8^1/_2$ day embryo. L-N are adjacent frontal sections from an $8^3/_4$ day embryo. A, E, H and L show bright-field images. B and I were probed with *Hox 1.6*. C, F, J and M were probed with *Hox 2.9*. D, G, K and N were probed with *Krox 20*. a, anterior; p, posterior; ne, neuroectoderm; m, mesoderm; pos, preotic sulcus; hb, hindbrain; fg, foregut pocket; h, heart; lpm, lateral-plate mesoderm; g VII- g VIII, 7th and 8th cranial ganglia; r3-r5, rhombomeres 3 to 5.


a
p
ne
pos
m
100µm
A
B
C
D
100µm
a
p
ne
m
hb
h
fg
E
s
F
G
a
hb
lpm
100µm
H
I
J
K
100µm
gVII
gVIII
r3
r4
r5
L
M
N
nc

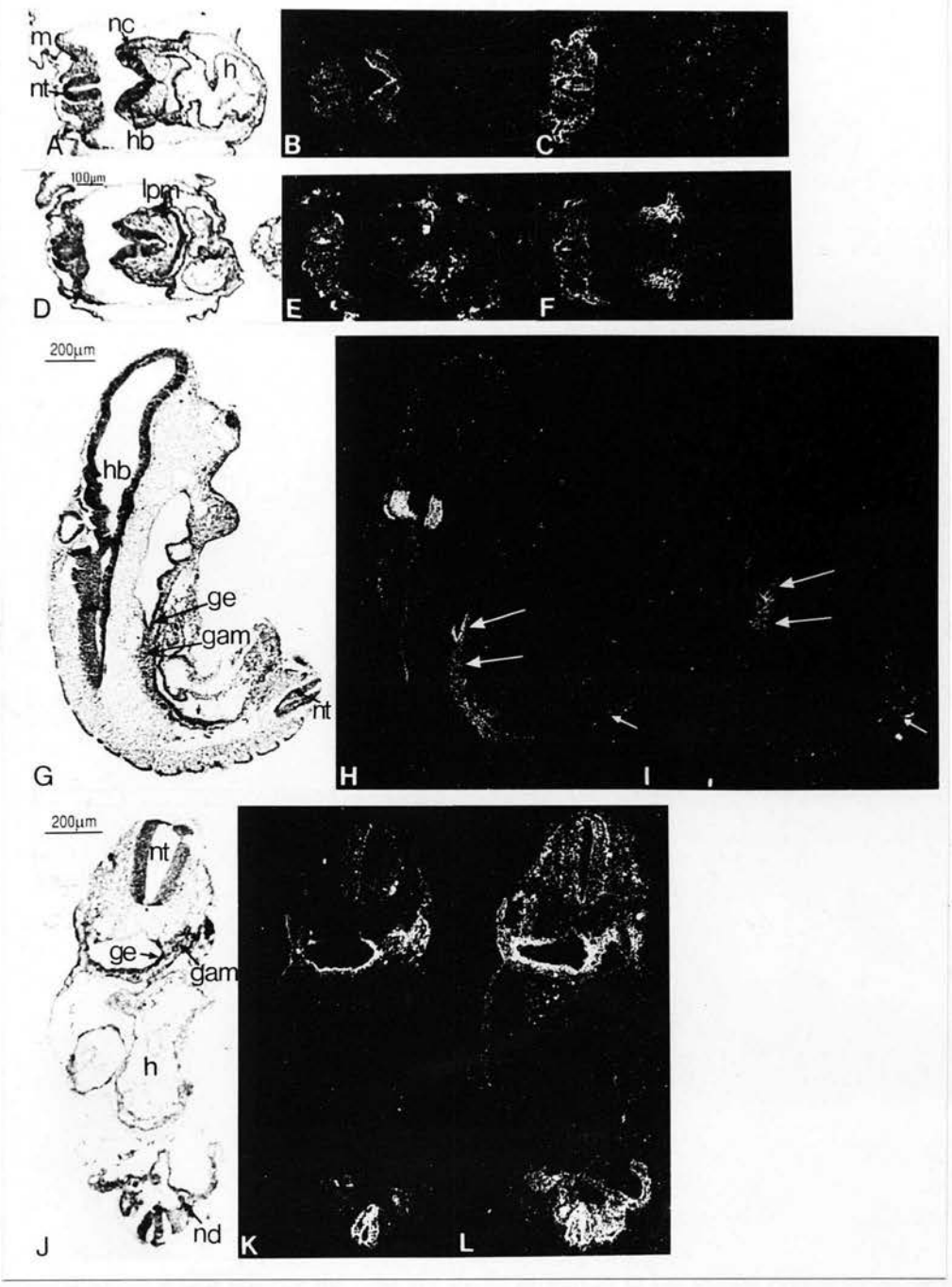


Figure 4.3. Expression of *Hox 2.9* and *Hox 1.6* at $8^3/_4$ and $9^1/_2$ days. A-F are transverse sections through an $8^3/_4$ day embryo. (A-C) sections at the level of rhombomere 4 in the hindbrain (on the right and the lower trunk (on the left); (D-F) sections cut at a more posterior level going through the lower hindbrain (on the right). B and E show dark-field images of A and D which were probed with *Hox 2.9*. C and F are adjacent sections to A and D respectively, and were probed with *Hox 1.6*. G-I are sagittal/frontal sections, and J-L are frontal sections, through a $9^1/_2$ day embryo. H and K were hybridised with *Hox 2.9* and I and L with *Hox 1.6*. m, mesoderm; nc, neural crest cells; nt, neural tube; hb, hindbrain; h, heart; lpm, lateral-plate mesoderm; se, surface ectoderm; ge, gut epithelium; gam, gut-associated mesoderm; nd, nephrogenic duct.

The rhombomeres, which are the morphological representation of segments within the hindbrain, are visible at $8^3/_4$ days (figure 4.2L-N). The rhombomeres are small and more evenly shaped at this stage than at later stages. We can now see that the expression domains of *Hox 2.9* and *Krox 20* within the developing hindbrain are perfectly coincident with rhombomere 4 in the case of *Hox 2.9* and rhombomeres 3 and 5 in the case of *Krox 20*. The expression boundaries have also become more sharply defined. These results show that the expression of *Hox 2.9* and *Krox 20* become localised within the hindbrain in an anterior to posterior order up to 6 hours before segments are visible.

At $8^3/_4$ days, labelling with *Hox 2.9* is also detected in the mesoderm lateral to rhombomere 4 in the region where sensory ganglia are condensing (figure 4.2L+M). Migrating neural crest cells which originate from rhombomere 4 also express *Hox 2.9* (figure 4.3A-B), and it is therefore likely to be the neural crest cell component of the ganglia that are positively labelled.

4.2.2 Hox 2.9 *and* Hox 1.6 *expression between 9 and 11 days.*

Between 9 and 10 days of development *Hox 2.9* and *Hox 1.6* are expressed within the neural tube in posterior regions only, in a way that is consistent with the expression domains retreating posteriorly, since there is a posterior to anterior gradient (figure 4.3J-L). This may relate to the process of maturation in the neural tube. A dorsoventral gradient of *Hox 2.9* expression within the neural tube is also visible (figure 4.4) and this relates to a period of cytodifferentiation in which sensory neurons are being produced in the dorsal region of the neural tube where *Hox 2.9* is most abundantly expressed. Dorsoventral sublocalisation of homeobox gene expression within the neural tube has previously been described (Bogarrad *et al.*, 1989). *Hox 2.9* is expressed most heavily within rhombomere 4 of the hindbrain (figure 4.3G+H). We have previously described how sharply defined this domain is at the cellular level (Murphy *et al.*, 1989). A series of sections through a 10 day embryo shows that a very narrow band of cells in the floor

plate of rhombomere 4 does not express *Hox 2.9* (figure 4.5), this compliments the fact that rhombomere boundaries do not extend into the floor plate.

Within the mesoderm, expression of both genes is now seen in gut-associated mesoderm at and below the level of the heart and in remaining pre-somitic mesoderm in posterior regions (figure 4.3G-L). In addition, *Hox 2.9* is expressed in the nephrogenic duct of the developing kidney (figure 4.3K). The domains of the two genes in the gut-associated mesoderm have the same AP restrictions, although *Hox 1.6* expression appears to be more extensive laterally, but this may simply reflect differences in the efficiencies of the two probes. There is also expression in the surface ectoderm adjacent to the labelled gut-associated mesoderm (figure 4.6A-B and 4.3J-L). Both genes are expressed in gut epithelium at the level of the fore limb bud (figure 4.3G-I). This is a derivitive of the endoderm and is therefore one of the few examples of endodermal expression of homeobox containing genes (Holland & Hogan, 1988b; Duprey *et al.*, 1988).

By $10^1/_2$ days the mesodermal expression of *Hox 2.9* and *Hox 1.6* has been down regulated (figure 4.6C-E) and is not detectable at $11^1/_2$ days (figure 4.6F-H). At $10^1/_2$ days *Krox 20* is no longer detectable in the hindbrain but *Hox 2.9* expression persists at a reduced level until $11^1/_2$ days (figure 4.6G). By $12^1/_2$ days no expression of *Hox 2.9* or *Hox 1.6* is detectable in the embryo.

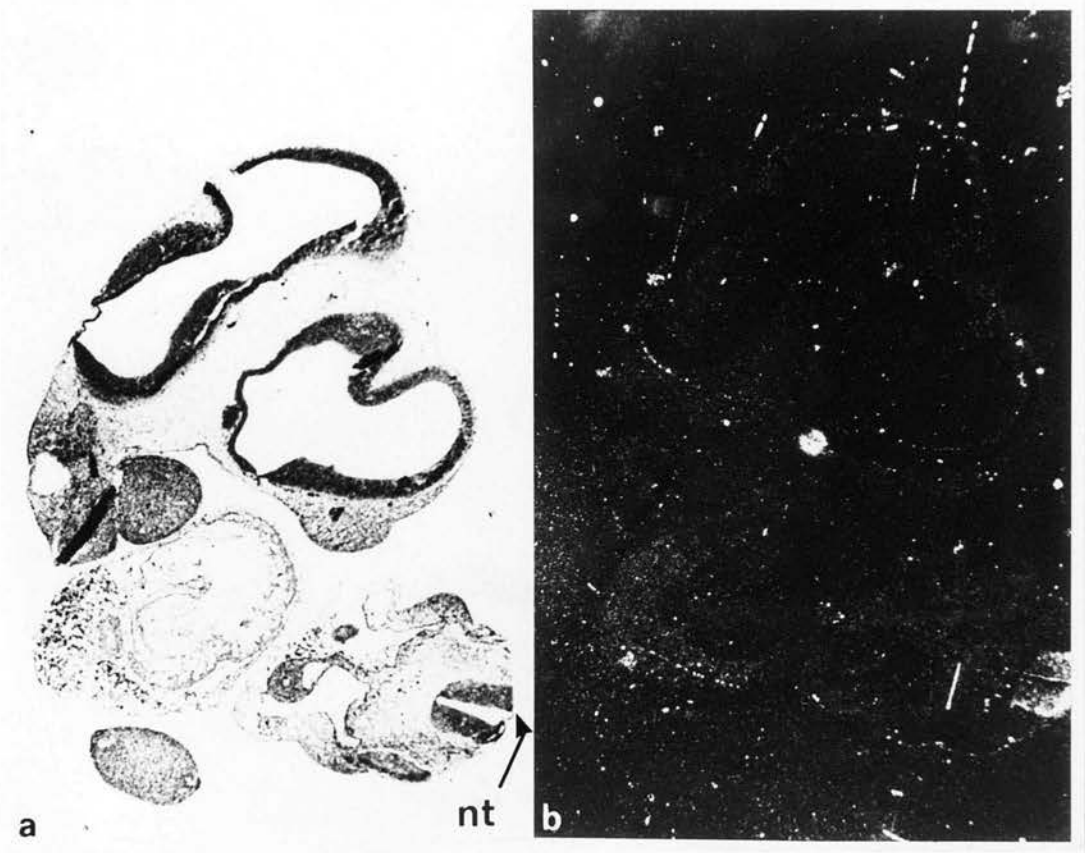


Figure 4.4. A section of a $9^1/_2$ day embryo, parasagittal in the head region and transverse in the trunk. The section was probed with Hox 2.9. The dark-field image in b shows expression restricted to the dorsal side of the trunk neural tube. nt, neural tube.

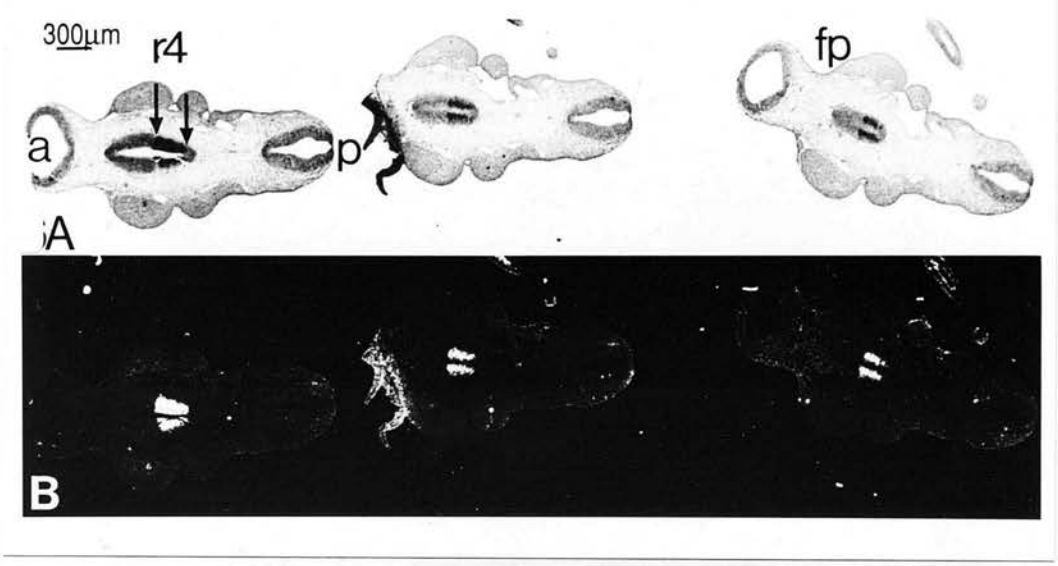


Figure 4.5. A series of adjacent frontal sections through a 10 day embryo, probed with *Hox 2.9*. The sections to the right are progressively more ventral and include the floor plate of the hindbrain which is not labelled with *Hox 2.9*. a, anterior; p, posterior; r4, rhombomere 4; fp, floor plate.

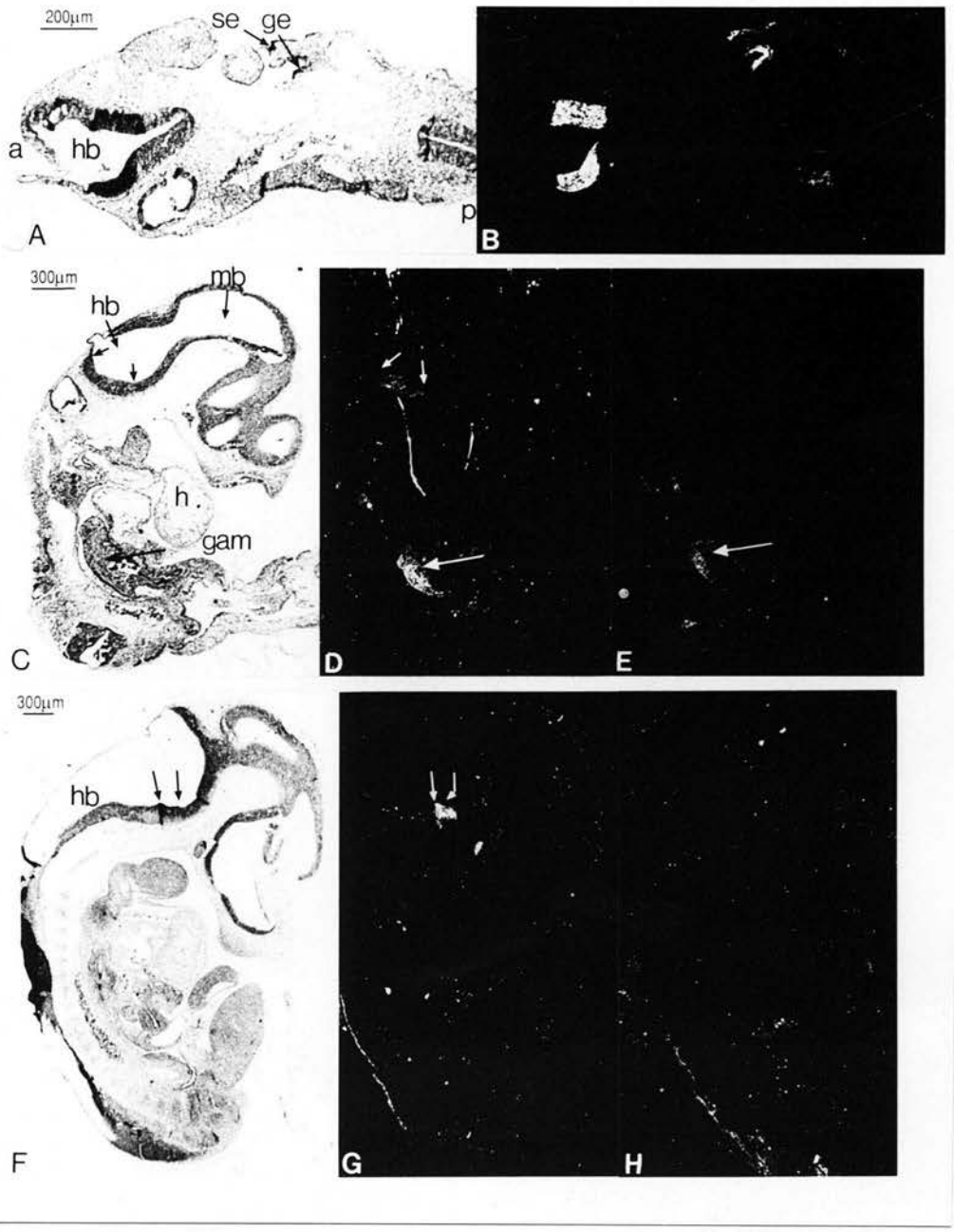


Figure 4.6. Expression of *Hox 2.9* and *Hox 1.6* between $9^1/_2$ and $11^1/_2$ days of development. (A-B) A frontal section through a $9^1/_2$ day embryo probed with *Hox 2.9*. (C-D) Adjacent sagittal sections from a $10^1/_2$ day embryo probed with *Hox 2.9* (D) and *Hox 1.6* (E). (F-H) Adjacent sagittal sections from an $11^1/_2$ day embryo probed with *Hox 2.9* (G) and *Hox 1.6* (H). se, surface ectoderm; ge, gut epithelium; hb, hindbrain; mb, midbrain; a, anterior; p, posterior; h, heart; gam, gut-associated mesoderm.

4.2.3 Expression of the differential transcripts of Hox 1.6.

Hox 1.6 is differentially transcribed, mature transcripts differing by a 203bp region 5' of the homeobox (LaRosa & Gudas, 1988b). Transcripts which contain this region code for a full length protein with a homeodomain, transcripts which lack this region code for a truncated protein with no homeodomain. In F9 teratocarcinoma cells both transcripts are produced, the relative amount of shorter transcript increasing from 10% to as high as 56% after treatment with retinoic acid which induces the cells to differentiate (LaRosa & Gudas, 1988b). We isolated 17 different *Hox 1.6* cDNA clones from an $8^1/_2$ day cDNA library and found that 7 did not contain the differentially spliced region, showing that both forms of transcript are produced in the early embryo (section 3.2.1).

The distribution of the *Hox 1.6* transcripts was investigated by *in situ* hybridisation. Figure 4.8 shows a diagramatic representation of the differentially spliced *Hox 1.6* transcripts and the strategy used to prepare probes capable of distinguishing them. *Hox 1.6* 3' is the probe that was used in the previously described experiments and is capable of detecting both types of transcript. Although this probe contains the homeobox region, the detection of an expression pattern distinct from that of *Hox 2.9*, its most closely related gene, indicates that it is specific for *Hox 1.6* transcripts. The stretch of 203bp that is only present in transcripts that code for a full length protein, was amplified by the polymerase chain reaction (PCR) method using oligos with extended non complimentary 5' ends that contain specific restriction enzyme sites (figure 4.7). In this way artificial *Hind* III and *Eco* RI sites were attatched respectively to the 3' and 5' ends of the diferentially spliced region. These sites were then used to directionally clone the amplified fragment into Bluescribe. Transcription *in vitro* from the T7 promoter produces an antisense probe capable of detecting only full length transcripts. Consecutive embryonic sections were hybridised with alternate *Hox 1.6* probes.

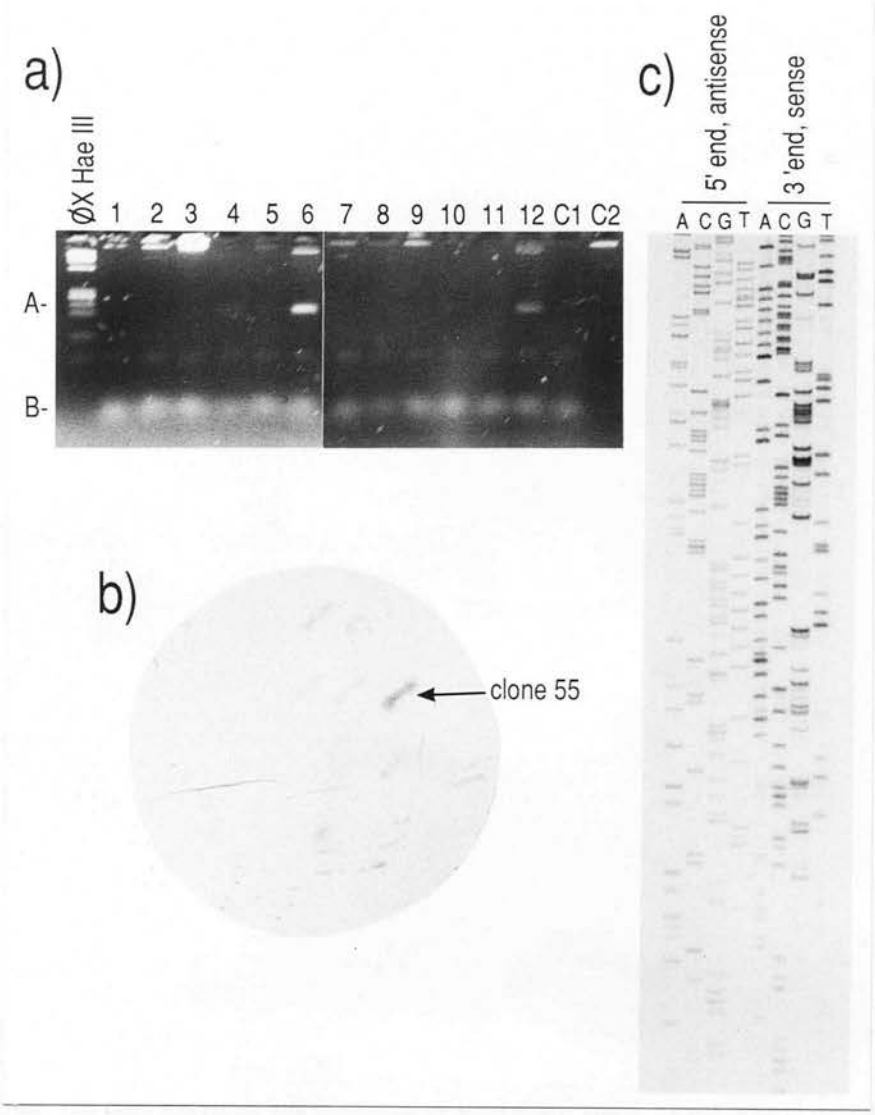


Figure 4.7. Cloning of the differentially spliced region of *Hox 1.6* (see section 2.2.5). a); an agarose gel showing the amplification of the region by PCR. Lanes 1-12 contain PCR reaction products with variable proportions of a number of template DNAs (lambda clones containing *Hox 1.6* cDNAs) and oligonucleotide primers. Lanes C1 and C2 contain control reactions in which only oligonucleotide (C1) or template DNA (C2) was added. Lanes 6 and 12 contain reactions in which an approximately 200bp region was successfully amplified. Using the artificial restriction sites at the ends of the amplified fragment, it was cloned into the vector Bluescribe. b); *E. coli* clones transformed with Bluescribe plasmid into which the amplified DNA in lane 6 of a) was ligated. These were probed with an oligonucleotide specific to the differentially spliced region of *Hox 1.6*. Clone 55 hybridised strongly. c); Sequencing reaction products, primed in both directions into the insert of clone 55, separated on an acrylamide gel and autoradiographed (section 2.9). This confirms that the insert contained in clone 55 is the differentially spliced region of *Hox 1.6*.

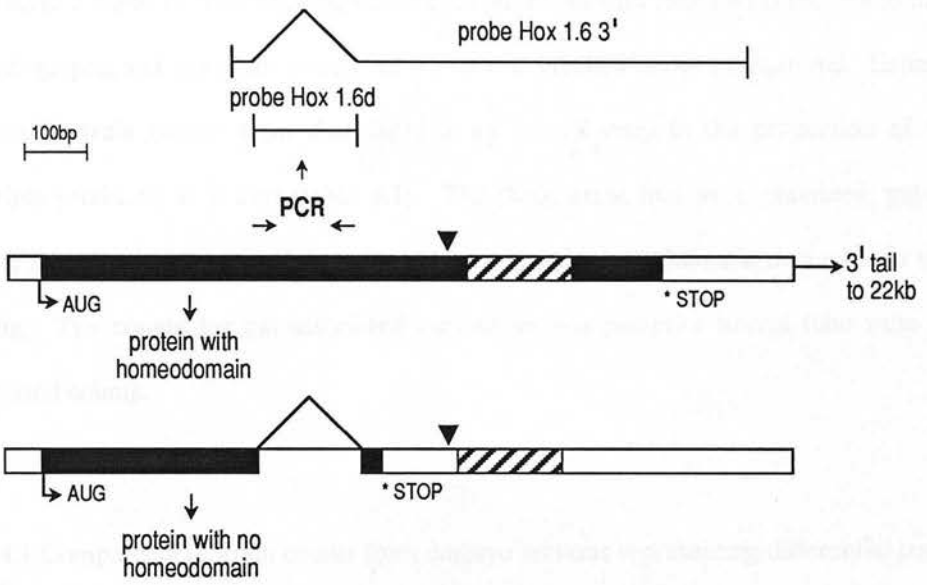


Figure 4.8 A representation of the differentially spliced products of the *Hox 1.6* gene. Solid shading represents open reading frames. Diagonal hatching indicates the homeobox sequence. The common splice site is marked by solid arrow heads. Single lines at the top of the diagram indicate the probes used for *in situi* hybridisation; *Hox 1.6* 3' can detect both transcripts whereas *Hox 1.6* d can detect only full length transcripts. *Hox 1.6* d was cloned following PCR amplification primed by synthetic oligonucleotides (represented by horizontal arrows).

At 8 days of development both *Hox 1.6* probes detected the same widespread domain of expression (figure 4.9A-C). The labelling with *Hox 1.6.d* was at a lower level (53-65%, table 1) but it shows that the full length transcript is being produced in the embryo at 8 days. At 9 days of development however, when the *Hox 1.6 3'* probe detects transcripts in a broad region of gut-associated mesoderm and gut epithelium, pre-somitic mesoderm, and posterior neural tube, the full length transcript is only detectable above background with *Hox 1.6.d* in the gut epithelium (figure 4.9D-I, table 1). The labelling of the gut epithelium with *Hox 1.6d* is too low to be visible in the photographs, but the grain counts show that it is labelled above background. Estimates from direct silver grain counts show that there is an overall drop in the proportion of full length transcripts produced at 9 days (table 4.1). The three areas that were examined; gut-associated mesoderm, gut epithelium and posterior neural tube, all show dramatic decreases in the level of labelling. The counts for gut-associated mesoderm and posterior neural tube were not above background counts.

Table 4.1 Comparison of grain counts from embryo sections representing differential transcripts of *Hox 1.6*.
Silver grain counts from *in situ* hybridised embryo sections and the ratio of mean counts with two *Hox 1.6* probes. ^{O} denotes the probe which detects both differential transcripts of *Hox 1.6*. ^{OO} denotes the probe that detects only full-length transcripts. The differences between the mean counts for the two probes were found to be statistically significant at less than the 1% level in all cases. With the exception of the values marked (*) all were significantly above background estimations.

	Hox 1.6 3' probe^{O}	Hox 1.6d^{OO}	Hox 1.6d/Hox 1.6 3'
8 day neuroectoderm	91.3 +/- 2.9	59.1 +/- 8.4	0.65
8 day mesoderm	84.1 +/- 5.1	44.6 +/- 4.3	0.53
9 day gut epithelium	130.0 +/- 6.9	26.1 +/- 12.2	0.20
9 day neural tube	60.3 +/- 4.9	4.5 +/- 2.7*	0.08
9 day mesoderm	40.2 +/- 10.6	4.3 +/- 4.6*	0.10

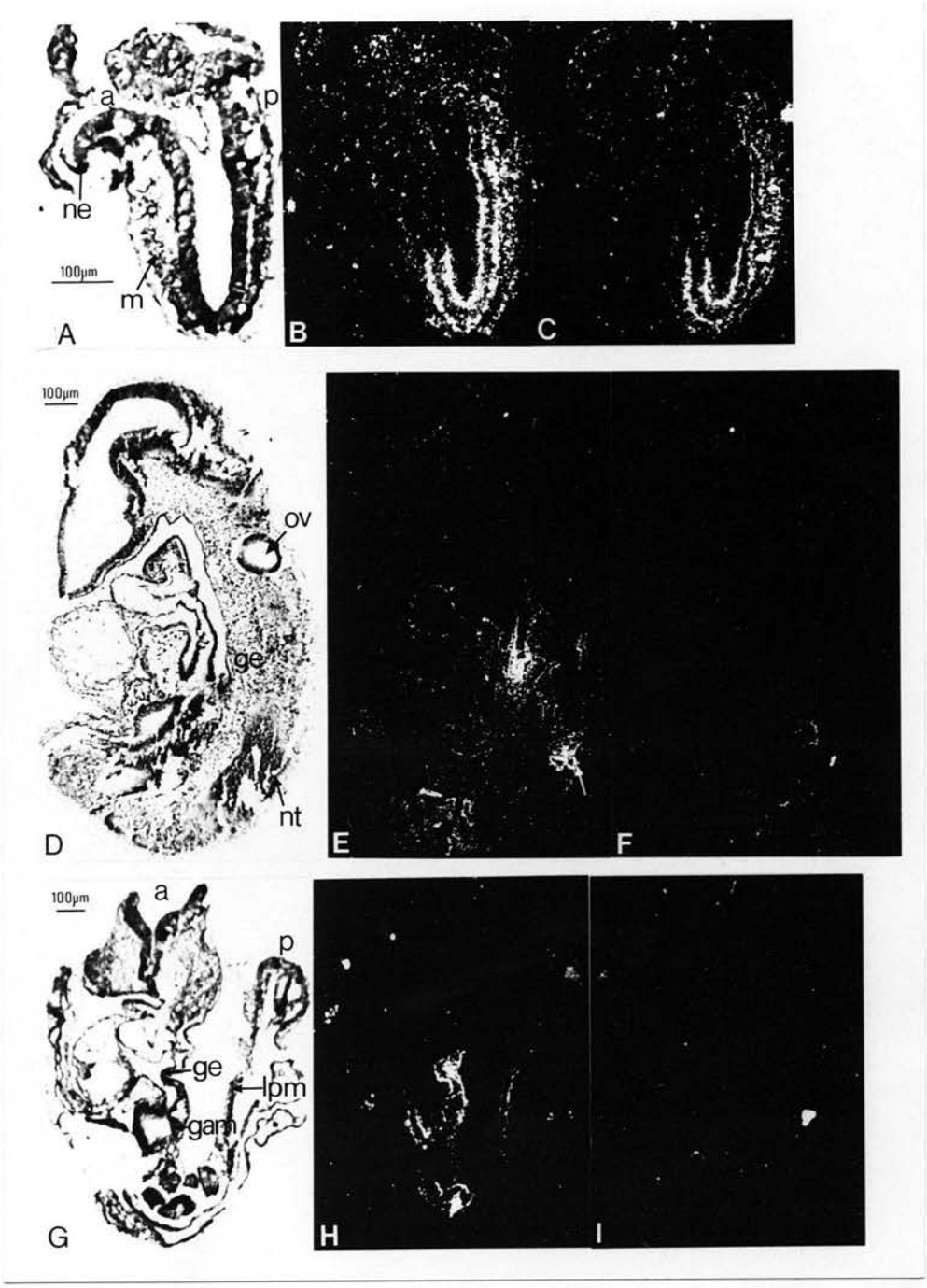


Figure 4.9. Expression of the differential transcripts of *Hox 1.6* at 8 and 9 days of development. (A-C) Sagittal sections through an 8 day embryo. (D-F) Sagittal sections, and (G-I) frontal sections through a 9 day embryo. B, E and H were probed with a fragment from the 3' end of *Hox 1.6* which detects both transcripts. C, F and I were probed with *Hox 1.6d* which detects only the full length transcript. The arrowhead denotes the anterior boundary of expression; a, anterior; p, posterior; ne, neuralectoderm; m, mesoderm; ov, otic vesicle; ge, gut epithelium; nt, neural tube; gam, gut-associated mesoderm; lpm, lateral-plate mesoderm.

4.3 Discussion.

4.3.1 A number of general features distinguish labial*-like genes from other* Hox *genes in the mouse.*

An interesting general feature of the expression of clustered homeobox genes, which is shared by vertebrates and *Drosophila*, is that position within the cluster is reflected in position along the body axis at which the gene is expressed (Akam, 1987; Scott & Carroll, 1987; Harding *et al.*, 1985; Graham *et al.*, 1989; Duboule & Dolle, 1989). In this respect *Hox 2.9* represents a special case in that it is positioned at the end, termed the 3' end, of the cluster, but the neighbouring gene to the 5' side, *Hox 2.8*, is expressed more anteriorly (Wilkinson *et al.*, 1989b). *Hox 2.8* has no equivalent gene in the Hox 1 cluster and so *Hox 1.6* is the most anteriorly expressed (Duboule & Dolle, 1989). Furthermore both mouse *labial*-like genes have exceptional expression patterns within the hindbrain at 9 days. The expression of *Hox 1.6* in the hindbrain is more transient than that of the other homeobox-containing genes in that no expression is detectable at 9 days. *Hox 2.9* expression disrupts the pattern observed with other *Hox 2* cluster members of sequential genes posessing anterior boundaries at two segment intervals (Wilkinson *et al.*, 1989b). *Hox 2.9* is the only *Hox 2* cluster gene to be uniquely expressed in a single rhombomere. Unlike other *Hox* genes which are generally expressed in overlapping domains in the somites and later the prevertebral column (Holland & Hogan, 1988b; Graham *et al.*, 1989; Duboule & Dolle, 1989), mouse *labial*-like genes are only expressed in the pre-somitic mesoderm with expression decreasing as somites condense.

4.3.2 Hox 2.9 *and* Hox 1.6 *expression: temporal and spatial similarities.*

Hox 2.9 and *Hox 1.6* share several features of their temporal and spatial expression patterns. Expression of both genes is initiated at the posterior end of the $7^1/_2$ day gastrulating embryo. By 8 days the expression patterns are indistinguishable with both genes occupying

domains which extend from the primitive streak into the presumptive hindbrain, with identical anterior boundaries within the neuroectoderm at the pre-otic sulcus. At a relatively early stage (before $8^1/_2$ days), the neural tube expression retreats along the AP axis so that at $9^1/_2$ days only very posterior regions of the neural tube express *Hox 2.9* and *Hox 1.6*, with the additional persistent expression of *Hox 2.9* in a distinct region of the hindbrain. There are also parallels in the mesodermal and ectodermal expression of *Hox 2.9* and *Hox 1.6*; in the pre-somitic and lateral plate mesoderm at $8^1/_2$ days where transverse sections reveal that they have the same AP limits; in the modified mesodermal expression at $9^1/_2$ days when both genes are expressed in gut-associated mesoderm; and in the lack of detectable mesodermal expression after 10 days. The timing, the extent, and the transient nature of the expression of these two genes indicate that they are responding to the same or similar signals in the embryo. The rapid loss of *Hox 2.9* and *Hox 1.6* RNA from the neural tube is either due to a loss of this stimulating signal or to simultaneous active repression of the genes. It also indicates that the transcripts have a rapid turnover rate, which has been shown for *Drosophila ftz* RNA with a half life of 6-8 minutes (Edgar *et al.*, 1986). The persistance of *Hox 2.9* expression in the hindbrain indicates that it can respond to an additional specific signal related to the segmentation of the hindbrain. This additional response leads to an intensification of the level of *Hox 2.9* in a subsection of the original domain. For this reason it will be important to investigate the control regions of these genes and to compare the binding sites for regulators that are present.

The unique expression of *Hox 2.9* within rhombomere 4 must represent a specialised function for *Hox 2.9*. It is possible that the earlier more widespread domain, from which this domain is derived and which is shared with *Hox 1.6*, is only functional in priming the later restricted expression. This is reminescent of *Drosophila* homeotic genes which have early widespread expression domains that become restricted to the corresponding functional domains (for review, Akam, 1987). Alternatively the broad expression of *Hox 2.9* and *Hox 1.6* at 8 days may be involved in positional signalling that is important prior to segmentation.

4.3.3 Homeobox genes are coordinately expressed in different germ layers early in development.

During the early phase of *labial*-like gene expression ($7^1/_2$ - $8^1/_2$ days) there appears to be co-ordinant expression in the ectodermal and mesodermal tissue layers resulting in corresponding AP limits in these two tissues. As development proceeds and the complexity of the embryo increases expression in both tissue layers becomes modified and there is little correspondence between the two layers. We therefore suggested that in the early stages of development basic AP positional domains are being similarly defined in the embryo as a whole, whereas in the later embryo developmental fields become more independent (Murphy & Hill, 1991).

Co-ordinate expression in the mesoderm and ectoderm early in development has previously been reported for other *Hox* genes in the mouse (*Hox 1.5*, Gaunt *et al.*, 1988) and the frog (*XlHbox 1,* Olivier *et al.*, 1988). It was suggested by DeRobertis *et al.* (1989) that this is established by expression in the mesoderm inducing expression in the overlying ectoderm since it is known from transplantation experiments in the amphibians that the mesoderm conveys AP positional information to the ectoderm (reviewed in Hamburger, 1988). An independent analysis of the expression of *Hox 2.9* by Frohman *et al.* (1990) has been interpreted by them as indicating that this is in fact the case for *Hox 2.9*. This is supported by their additional observation that *Hox 2.9* is expressed in mesoderm without expression in overlying ectoderm at $7^1/_2$ days, prior to co-ordinate expression. The $7^1/_2$ day embryos analysed by us appeared to show co-ordinate expression in mesoderm and ectoderm, however this may be due to slight differences in the stages of the embryos in the two studies. The alternative to mesodermal induction of ectodermal expression is that expression in both layers is co-ordinately controlled by common AP positional determinants. As was pointed out by DeRobertis *et al* (1989) it is unlikely that such a determinant is a diffusible morphogen since diffusion constants would be very different in mesoderm and ectoderm. However another unknown common mechanism may be responsible. Our results do not distinguish between mesodermal induction of ectoderm or co-ordinate induction of the two layers, but they do suggest the existence of basic AP positional domains that include more than one tissue layer early in development.

4.3.4 The relationship between the expression of Hox 2.9 *and hindbrain segmentation.*

The localised expression of *Hox 2.9* and *Krox 20* within specific rhombomeres has previously been described (Wilkinson *et al.*, 1989a; Murphy *et al.*, 1989). The analysis presented here focuses on earlier *Hox 2.9* expression and details further the role this gene plays in hindbrain segmentation. Furthermore, together with *Krox 20,* these genes provide useful molecular markers in studying the process of segmentation in the hindbrain. *Krox 20* is first expressed in two domains within the hindbrain which will become rhombomeres 3 and 5. This expression is first initiated in the more anterior domain, followed by initiation in the more posterior domain with both domains expressing *Krox 20* prior to the appearance of rhombomeres (Wilkinson *et al.*, 1989a). The results presented here (summarised in figure 4.10) show that *Hox 2.9* is expressed in the hindbrain, with a defined anterior boundary, before rhombomeres are visible, at the time that *Krox 20* is expressed in a single domain. This is supported by the findings of Frohman *et al.*, (1990) but contrary to those of Wilkinson *et al.* (1989b) who reported that *Hox 2.9* expression in the hindbrain is initiated in its restricted domain after both domains of *Krox 20* expression are established. We find that at approximately the time that *Krox 20* expression is initiated in the second domain, *Hox 2.9* expression becomes localised to the region of the hindbrain that will form rhombomere 4. At no time did we observe an overlap in the expression of *Hox 2.9* and *Krox 20* and once the domains are established they have sharp planar boundaries indicating that there is little or no cell mixing occurring between the domains. The expression pattern of these genes therefore indicate that compartmentaslisation of the hindbrain begins in the 8 day embryo and progresses in an anterior to posterior direction. By $8^1/_2$ days the segmental units represented by rhombomeres 3, 4 and 5 have been defined.

After rhombomeres are visible we show that in rhombomere 4 *Hox 2.9* is not expressed in the floorplate. Fraser *et al.* (1990) have demonstrated that there are no rostro-caudal cell lineage restrictions in the floor plate of the chick hindbrain and that the floor plate also lacks visible rhombomere boundaries. *Hox 2.9* is therefore only expressed in the part of rhombomere 4 that is obviously segmented. Rhombomeres are transient structures which disappear by day 12. The *Hox*

2.9 rhombomere 4 expression is not detectable after $11^1/_2$ days and therefore expression persists throughout the period that rhombomere 4 exists. This data further suggests that *Hox 2.9* is involved in specifying the identity of the developmental compartment defined as rhombomere 4.

Our description of how the *Hox 2.9* expression pattern evolves has given us insights into the timing and progression of hindbrain segmentation (Murphy & Hill, 1991). Antibodies have been prepared which recognise the closely related chicken *labial*-like gene, *Ghox lab*. These have been used in a similar investigation of the process in the chicken (Sundin & Eichele, 1990). It is very valuable to produce specific antibodies as only an investigation at this level can reveal where the mRNA is translated to protein (Gaul *et al.*, 1987) and where the protein is localised within the cell. The greater resolution possible in the detection of protein with antibodies and the added advantages in obtaining and handling chicken embryos, have allowed them to describe a number of aspects of the pattern in more detail. We noticed a progressive sharpening of the *Hox 2.9* expression boundaries and it was shown by Sundin & Eichele that before and shortly after rhombomere 3 boundaries appear, a small number of cells within rhombomere 4 express *Hox 2.9*. In our early analysis we showed the precision of the *Hox 2.9* expression boundary (figure 3.12) where it appeared that all cells within rhombomere 4 were labelled (Murphy *et al.*, 1989). The chicken study has revealed that this is the case. We showed that the posterior boundary of rhombomere 4 expression was defined when the more posterior *Hox 2.9* expression decreased. It appears that *Ghox lab* expression increases within the region of rhombomere 4 prior to a decrease in rhombomeres 5 and 6. It was also clear that in the chicken morphological boundaries appear before the rhombomeric expression domain is established. Clear rhombomeric boundaries were not visible in embryonic sections at the time that expression was localised in the mouse. This may reflect a real difference in the morphological appearance of segments in the mouse and the chicken. However the chicken analysis indicates that rhombomeres are defined prior to the localisation of *Ghox lab* expression.

It appears that the development of the hindbrain is under a complex regime of regulatory controls since the expression patterns of *Hox 2.9, Hox 1.6* and *Krox 20* within the developing

hindbrain follow different modes as segmentation occurs. Both *Hox 2.9* and *Hox 1.6* are expressed early in broad domains with sharp anterior boundaries within the hindbrain. *Hox 2.9* later becomes localised to a single segment at which time *Hox 1.6* is no longer expressed. *Hox 2.9* and *Krox 20* are expressed in a segmental pattern, however *Hox 2.9* results from the modification of a broad region of expression and *Krox 20* is initiated in distinct domains. Whereas the segmental expression of *Hox 2.9* and *Krox 20* appears to be established at the same time, *Hox 2.9* expression persists for a longer period. These genes will be important in understanding the positional signalling events and the regulatory elements involved in the process of hindbrain segmentation.

4.3.5 Prepatterning of neural crest cells in the hindbrain.

Hox 2.9 is also expressed in the sensory ganglia associated with rhombomere 4 and in the neural crest cells that migrate from rhombomere 4. It is not known if all the labelled neural crest cells are destined to contribute to the sensory ganglia or if some migrate further and contribute to the branchial arch structures. The expression of *Hox 2.9* specifically in the neural crest cells that arise from rhombomere 4 supports the idea of neural crest cells being patterned according to their rhombomeric origin (Couly & LeDouarin, 1990). Transplantation studies show that this is the case for cranial neural crest cells (Noden, 1983). Neural crest cells are known to follow specific migratory pathways maintaining the AP order in which they arise (Tan & Morriss-Kay, 1986; Noden, 1975). Patterning of neural crest cells according to rhombomeric origin would therefore extend the segmental unit to regions outside the neuroectoderm.

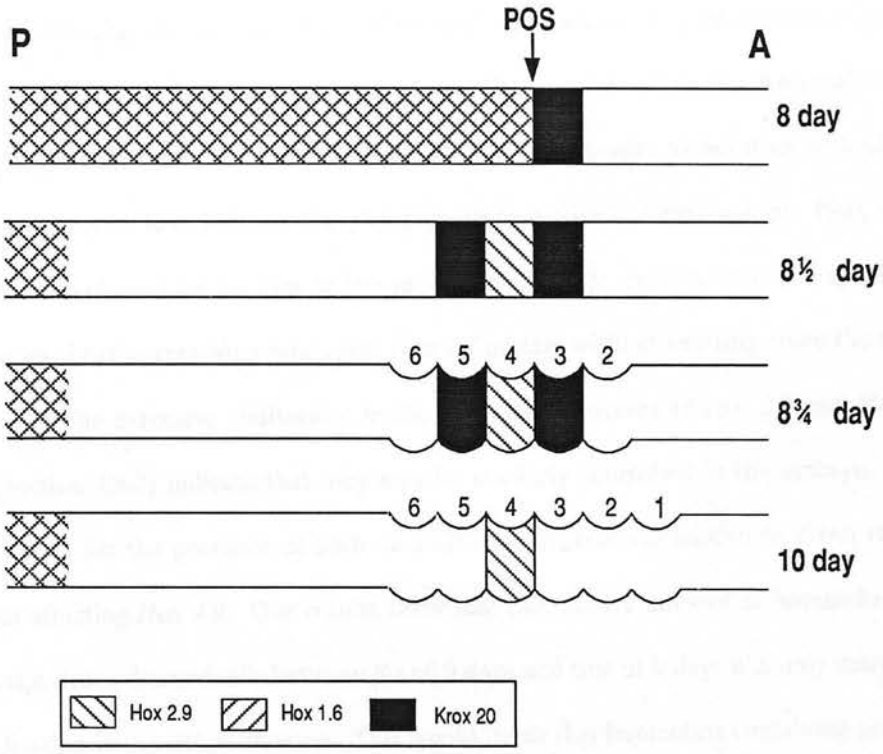


Figure 4.10 A diagramatic representation of the expression patterns of *Hox 2.9*, *Hox 1.6* and *Krox 20* in the CNS during segmentation of the hindbrain. The first appearance of the rhombomeres is represented at $8^{3}/_{4}$ days, at which time at least 5 rhombomeres are visible. The diagram does not represent the relative size of the CNS at various stages.

4.3.6 Hox 1.6 *differential transcripts; a change in their relative proportions as development proceeds.*

The alternate *Hox 1.6* transcripts produced in the embryo code for a homeodomain protein and a truncated non homeodomain protein. It is possible that the truncated protein has an independent function in the developing embryo or production of the spliced RNA may simply be a means of silencing the *Hox 1.6* direct DNA binding function. It is impossible at present to tell if the truncated protein is functional and if so what that function might be. Although it does not have the capacity to directly bind DNA it may maintain the capacity to interract with other regulators and in this way be involved in a complex regulatory network (Benezra *et al.*, 1990, for an example of such a regulatory mechanism in the myogenic system). Alternatively the splicing mechanism may be involved in removing functional *Hox 1.6* protein without shutting down the transcription of the gene. The extensive similarities in the expression patterns of *Hox 2.9* and *Hox 1.6* reported here (section 4.3.2) indicate that they may be similarly controlled in the embryo. This offers an explaination for the presence of such an additional control mechanism to down regulate *Hox 1.6* without affecting *Hox 2.9*. Our results show that the relative amount of homeodomain producing transcript drops dramatically between 8 and 9 days and that at 9 days it is only detectable by *in situ* hybridisation in the gut epithelium. This would imply that homeobox containing protein from both genes is required to pattern the 8 day embryo but there is a greater requirement for *Hox 2.9* as a transcription factor at later stages.

Chapter 5

The effect of excess retinoic acid on segmentation of the hindbrain.

5.1 Introduction.

The current evidence that retinoic acid (RA) acts as a morphogen in the developing embryo is very convincing (reviewed by Brockes 1989; Brockes 1990; section 1.2.6). The single fact that RA has the ability to respecify positional identity in the chick limb bud and, as such, mimic the effect of transplantation of the zone of polarising activity (ZPA) demonstrates its effect as a morphogen (Tickle *et al.*, 1982). In addition a long list of other findings support this theory and give us some clues as to the mechanisms that are involved (see section 1.2.6).

RA is normally acquired by the mammalian embryo through metabolism of retinol (vitamin A) supplied via the maternal bloodstream. The endogenous level can therefore be regulated. The level of RA experienced by the developing embryo is critical, with extreme teratogenic effects resulting from abnormal levels (Morriss, 1972; Morriss & Thorogood, 1978; Lammer *et al.*, 1985; Webster *et al.*, 1986). When early rat embryos (8-10 days) are exposed to excess RA, either *in vitro* or *in vivo*, a number of craniofacial defects are observed (Morriss, 1972; Morriss & Thorogood, 1978). The CNS and the neural crest cell derivitives appear to be the major targets. In the days following treatment there is later neural fold closure and later or slower neural crest cell migration. The treated embryos have less cephalic mesoderm which may be responsible for delayed closure of the neural folds which lack normal support from underlying mesoderm (with exencephaly in extreme cases). The major defect in the early embryo (10 days) however, is a shortening of the neural plate at the anterior end, causing a rostral shift in the position of the otic vesicle relative to the pharyngeal arches, so that it lies opposite the first (mandibular) arch instead of the second (hyoid) arch (figure 5.1).

In the later embryo (20 days) bone formation in the maxillary and mandibular regions is disrupted. In the mandibular region, Meckels' cartillage is shortened and the bones of the inner ear are absent or incomplete. In the maxillary region an ectopic bone forms. The correct development of these regions depends on the migration of neural crest cells from the brain into the

pharyngeal arches and the maxillary process. Mapping of neural crest cell migration showed that cells from the hindbrain migrate into the pharyngeal arches and those from the midbrain into the maxillary process (Tan & Morriss-Kay, 1986). It has also been shown that cephalic neural crest cells, unlike those in the trunk, are patterned according to their position of origin within the CNS (Noden, 1983). With this in mind, cranial skeletal malformations can be interpreted as a result of inappropriate neural crest cell migration following the rostral shift in the brain relative to the ventral side of the embryo (Morriss & Thorogood, 1978). According to Morriss & Thorogood (1978) neural crest cells enter the maxillary region of treated embryos before they migrate from the midbrain of controls. The major craniofacial effects of RA may therefore be due to an alteration of the AP axis primarily in the ectodermal layer leading to a shift in the spatial relationship of the germ layers. Experiments on the effect of RA treatment on early *Xenopus laevis* embryos also show that the anterior end is most sensitive and there is evidence that anterior neural tissue is transformed into a more posterior specification (Durston *et al.*, 1989; Sive *et al.*, 1990) further indicating a role for RA in AP axis formation, at least in the head.

When mammalian embryos are exposed to RA later in development (11 days in the mouse) defects are observed in the developing limbs with a reduction or absence of the long bones (Kochar, 1973; Satre & Kochar, 1989). Therefore there are two separate periods of sensitivity to RA indicating that it is involved in separate systems of pattern formation in the embryo at different times; earlier in the patterning of the brain and face and later in the limb.

The aims of the experiments described here were two fold:

(1) RA is a potential morphogen, and nuclear receptors for RA probably act by regulating the expression of sets of responsive genes (section 1.2.6). In the *Drosophila* system, homeobox and zinc finger motif encoding genes respond to spatial information in the embryo, including morphogenetic gradients (section 1.3.5). Homeobox genes are also known to be induced by RA treatment of F9 culture cells (Colberg-Poley *et al.*, 1985a; Mavilio *et al.*, 1988; LaRosa & Gudas, 1988b; Simeone *et al.*, 1990). This was shown to be the case for *Hox 2.9*, and the closely related

gene *Hox 1.6* in section 3.2.5. *Hox 2.9* and the zinc finger gene *Krox 20* are therefore potential targets for RA in the embryo. This is a particularly interesting relationship to explore since *Hox 2.9* and *Krox 20* are themselves regulatory genes and are capable of defining finer levels of positional information. It is therefore important to examine the effect of excess RA on the expression of *Hox 2.9* and *Krox 20.*

(2) As described above, RA treatment leads to a shortening of the anterior neural plate. *Hox 2.9* and *Krox 20* are normally expressed in precise segmental domains of the developing mouse hindbrain reflecting the segmental organisation that underlies hindbrain development (Lumsden & Keynes, 1989; Murphy & Hill, 1991). This pattern was characterised in detail in chapter 4. This system therefore offers a unique opportunity to investigate the effect of excess RA on the segmentation of the hindbrain.

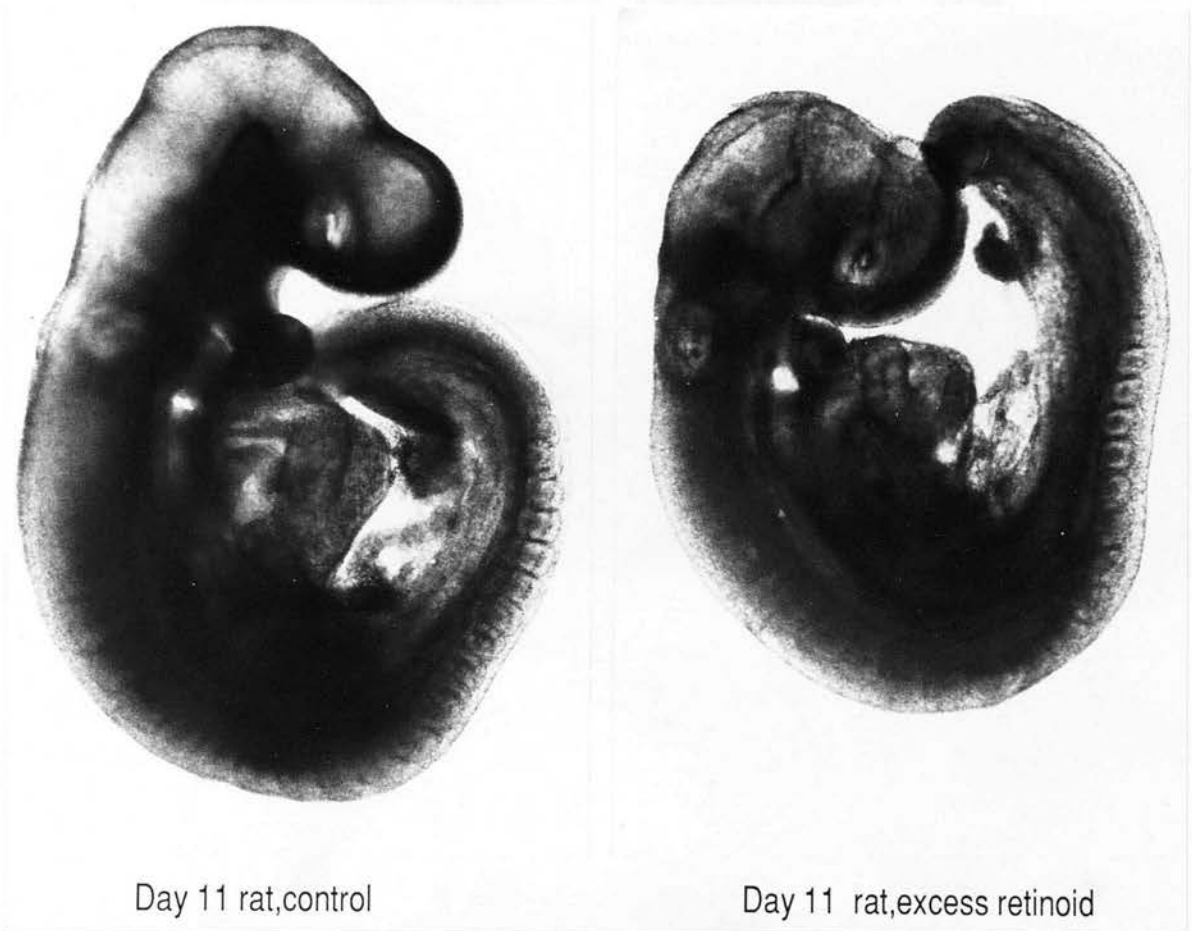


Figure 5.1. 11 day rat embryos showing the effect of treatment with excess RA at the neural fold stage of development. The treated embryo on the right clearly shows the rostral shift in the position of the otic vesicle from the position of the 2nd paryngeal arch (seen in the embryo on the left) to the position of the 1st paryngeal arch. (Photographs courtesy of G.M. Morriss-Kay).

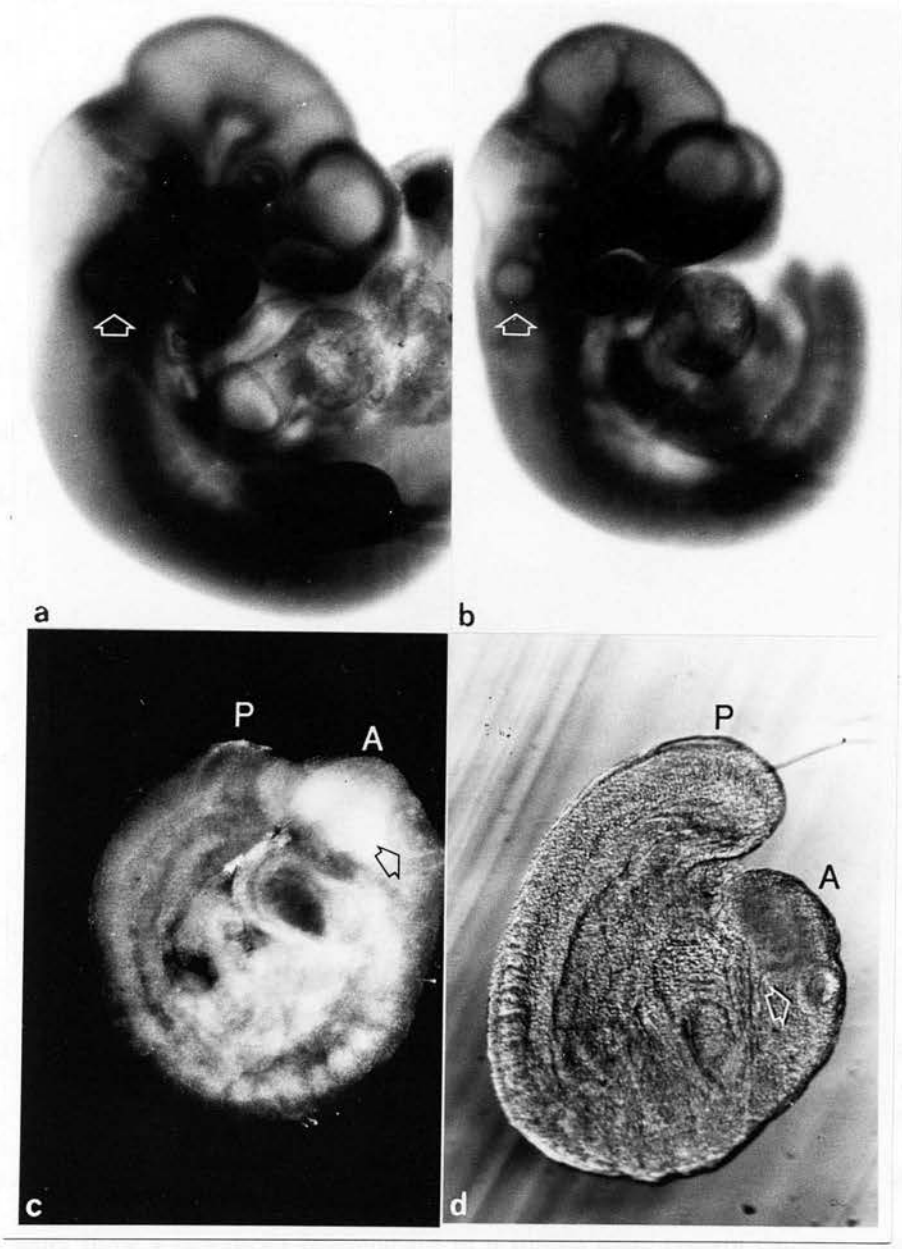


Figure 5.2. $9^1/_2$ day mouse embryos which had been exposed to excess RA at 8 days, showing variability in the severity of the effects. a and b are categorised as ot 1, with the otic vesicle shifted rostrally to the level of the 1st pharyngeal arch. c and d are more severly affected showing extreme reduction in the forebrain and midbrain and a more exaggerated rostral shift of the otic vesicle (categorised as ot max). A, anterior; P, posterior. The open arrow indicates the otic vesicle.

5.2 Results.

Female mice were dosed with RA on day 7 + 18 hrs or day 8 of pregnancy (see materials and methods, 2.10). The developing embryo is then at the neural fold stage and is still undergoing gastrulation. Somitogenesis has just begun and the segmental domains within the hindbrain are also just beginning to be defined (Murphy & Hill, 1991). The dose of 10 or 12mg / kg of all trans RA in arachis oil, was administered orally. The embryos were collected for examination at day 9, day 9 + 12hrs or day 10 of development. Since this type of treatment has variable effects on the developing embryo, the embryos recovered were categorised according to phenotype. Two categories were used for further examination: (1). Ot 1; where the otic vesicle is shifted from the level of the second pharyngeal arch to the level of the first pharyngeal arch; (2). Ot max; where the otic vesicle is shifted to the level of the maxillary process. Figure 5.2 shows examples of both categories. Control embryos were given 0.6ml of vehicle alone and were examined in parallel. This treatment did not produce any of the abnormalities observed following RA exposure. Maternal dosing and collection of embryonic material were carried out by Gillian Morriss-Kay at Oxford University.

5.2.1 The effect of excess RA on the morphology of the brain.

Examination of whole and sagitally halved embryos by scanning electron microscopy was carried out by G. Morriss-Kay (Morriss-Kay *et al.*, in preparation). This revealed that cranial neurulation is retarded in RA treated embryos, so that cranial neural folds of embryos with 11 or 12 somite pairs show a superficial resemblence to 9 somite stage control embryos. In sagitally halved 12 somite stage control embryos, up to 7 rhombomeric sulci of approximately equal size could be distinguished. In 12 somite ot 1 embryos there was an enlarged rhombomere-like structure in the position of rhombomere 1, just posterior to the midbrain. In some cases there was

a smaller sulcus caudal to this one but no others could be distinguished. In 16-20 somite stage ot max embryos, a large number of irregular sulci and gyri extended from the trunk region to the midbrain/forebrain junction. However, these were assymetrical and had the appearance of undulations rather than rhombomeres indicating that they do not reflect an underlying segmental organisation.

The forebrain and midbrain regions of RA treated embryos were consistently smaller than those of controls. The abnormally rostral position of the otic pit in RA treated embryos could clearly be seen externally in 11-16 somite stage embryos (figure 5.2).

Histological examination of embryo sections used for *in situ* hybridisation was carried out using bright field microscopy. At 9 days coronal and sagittal sections of control embryos clearly showed the characteristic rhombomeric structure of the hindbrain (figures 5.3 and 5.4). RA treated embryos showed the enlarged rhombomere-like structure described above (figure 5.4) while the remainder of the hindbrain lacked sulci and gyri. In ot max embryos at 10 days the neural tube and hindbrain showed irregular undulations (figure 5.). The comparative size of the forebrains of control and ot 1 embryos is shown in figure 5.

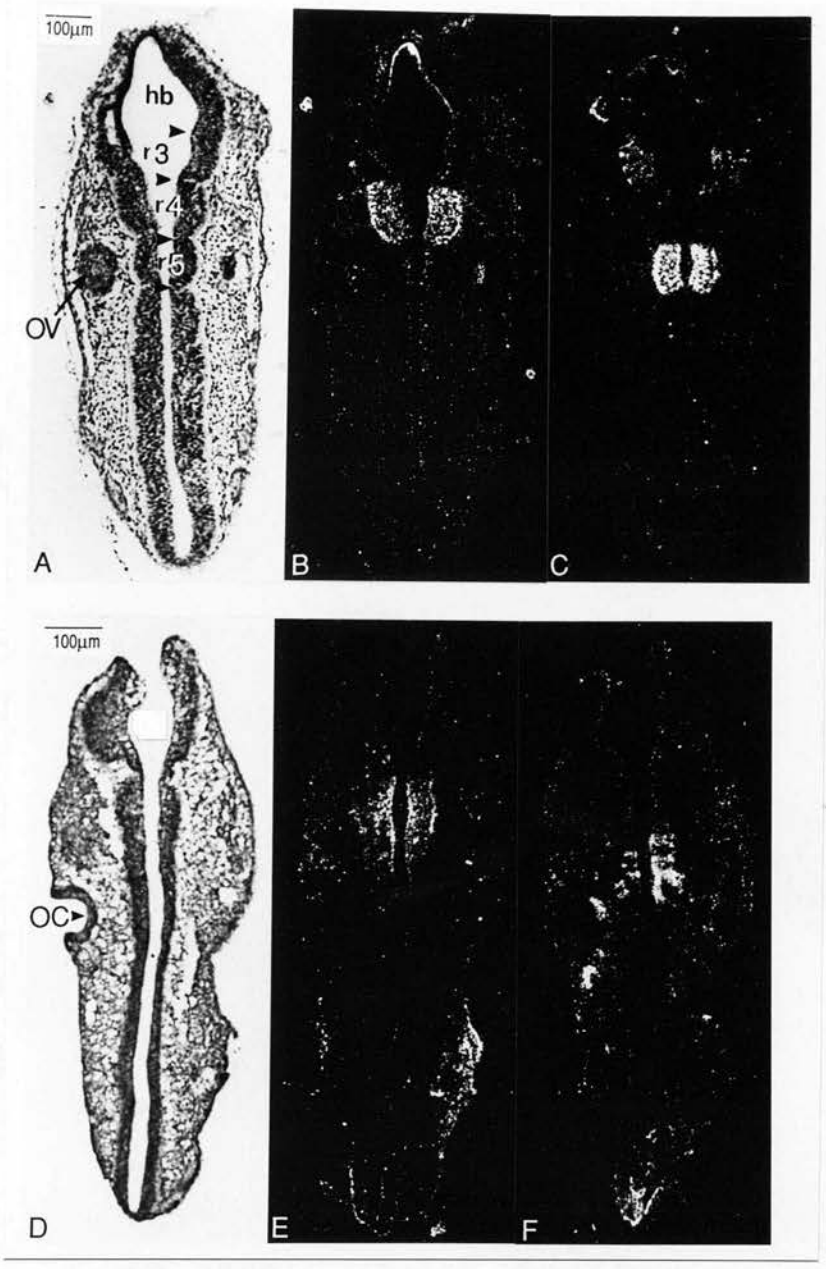


Figure 5.3. The effect of excess RA on the expression of *Hox 2.9* and *Krox 20* at $9^1/_2$ days. (A-C) Adjacent frontal sections from a control embryo showing the normal expression pattern of *Hox 2.9* (B) and *Krox 20* (C). (D-F) Adjacent frontal sections from a $9^1/_2$ day embryo which was exposed to excess RA at 8 days, showing a lack of morphological rhombomeres and the altered expression of *Hox 2.9* (E) and *Krox 20* (F). hb, hindbrain; r3-r5, rhombomeres 3 to 5; ov, otic vesicle; oc, otic cup.

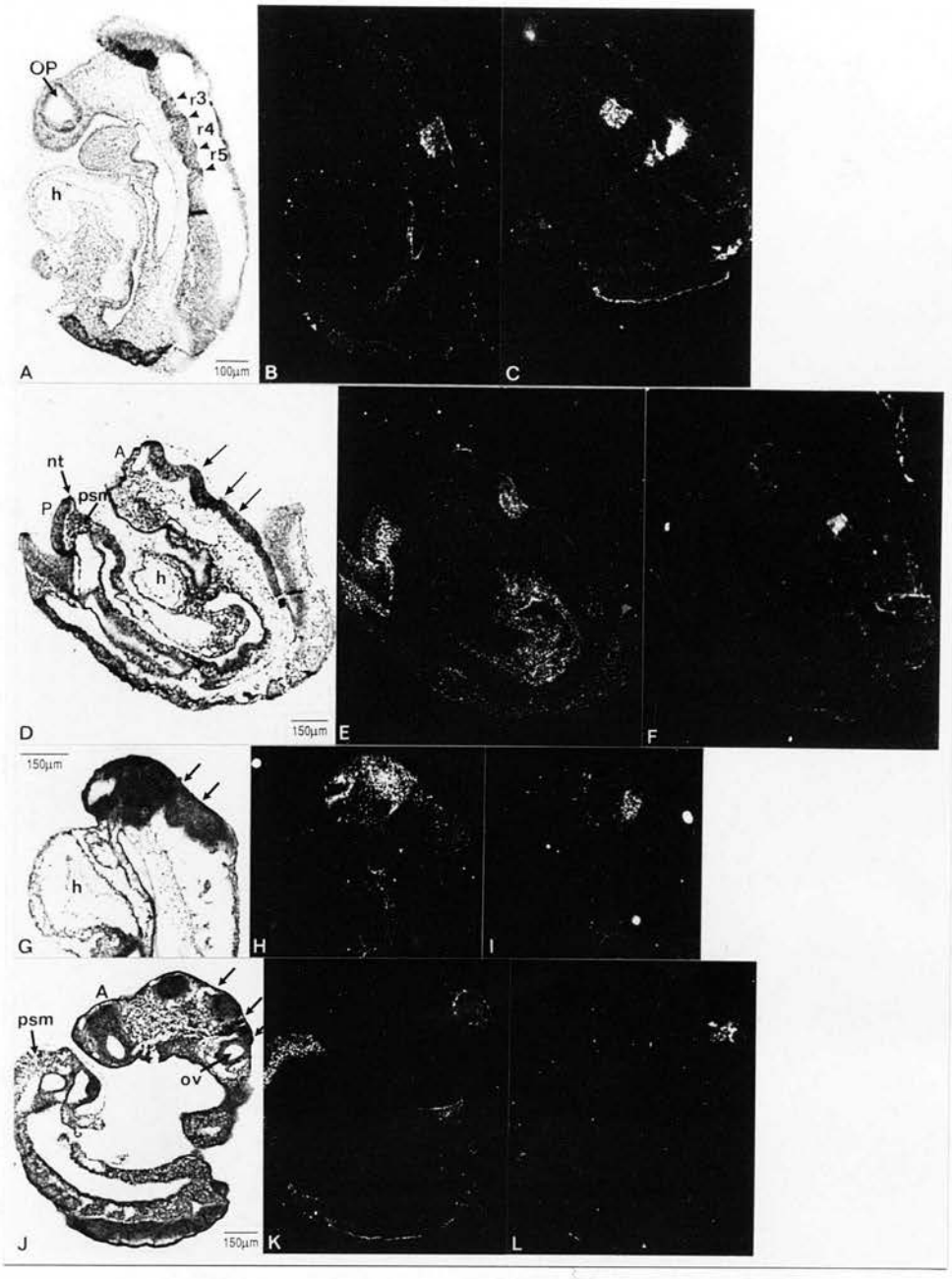


Figure 5.4. *Hox 2.9* and *Krox 20* expression in RA treated embryos at $9^1/_2$ days. (A-C) Control embryo sections showing expression of *Hox 2.9* (B) and *Krox 20* (C). (D-F) An ot 1 embryo showing expression of *Hox 2.9* (E) in an enlarged rhombomere-like structure just posterior to the hindbrain/midbrain junction and expression of *Krox 20* (F) only posterior to the *Hox 2.9* domain. (G-I) ot 1 embryo sections probed with *Hox 2.9* (H) and *Krox 20* (I). (J-L) Sections from an ot max embryo probed with *Hox 2.9* (K) and *Krox 20* (L) show the relationship between the expression domains and the otic vesicle. op, optic vesicle; r3-r5, rhombomeres 3 to 5; h, heart; A, anterior; P, posterior; nt, neural tube; psm, presomitic mesoderm; ov, otic vesicle.

5.2.2. The expression of Hox 2.9 *and* Krox 20 *in the hindbrain is shifted anteriorly following treatment with excess RA.*

Paraffin-embedded sections of treated and control embryos were hybridised *in situ* with *Hox 2.9* and *Krox 20* probes (described in section 2.7.5). The expression patterns of these genes in control embryos were found to be as previously described (chapter 4); in the hindbrain at day 9 *Hox 2.9* is expressed in rhombomere 4 and *Krox 20* in rhombomeres 3 and 5. In ot 1 embryos at 9 days, *Hox 2.9* is expressed in a domain which coincides with an enlarged rhombomere-like structure anterior to the otic vesicle (figure 5.3, 5.4). This domain is in an approximately equivalent position to rhombomere 4 in relation to the otic vesicle. However, the otic vesicle is shifted anteriorly with respect to the pharyngeal arches. Morphological analysis of treated embryos (section 5.2.2) revealed that this rhombomere-like structure is anteriorly adjacent to the midbrain, in the normal position of rhombomere 1 (Morriss-Kay *et al.*, in preparation). In 9 day ot 1 embryos, *Krox 20* is expressed in a single domain in the hindbrain, whereas in control embryos at the same stage, it is expressed in two domains. No expression of *Krox 20* is detected anterior to the *Hox 2.9* expression domain in a position equivalent to rhombomere 3, which normally expresses *Krox 20*. The single domain of expression of *Krox 20* lies posterior to *Hox 2.9* expression, in the region equivalent to rhombomere 5. In control embryos, rhombomere 5 lies directly adjacent to the otic vesicle but the domain of expression of *Krox 20* in treated embryos extends slightly more anteriorly. Therefore the otic vesicle and the domains of expression of *Hox 2.9* and *Krox 20* are shifted anteriorly with respect to the ventral side of the head, and furthermore the expression domains do not maintain their precise spatial relationship to the otic vessicle.

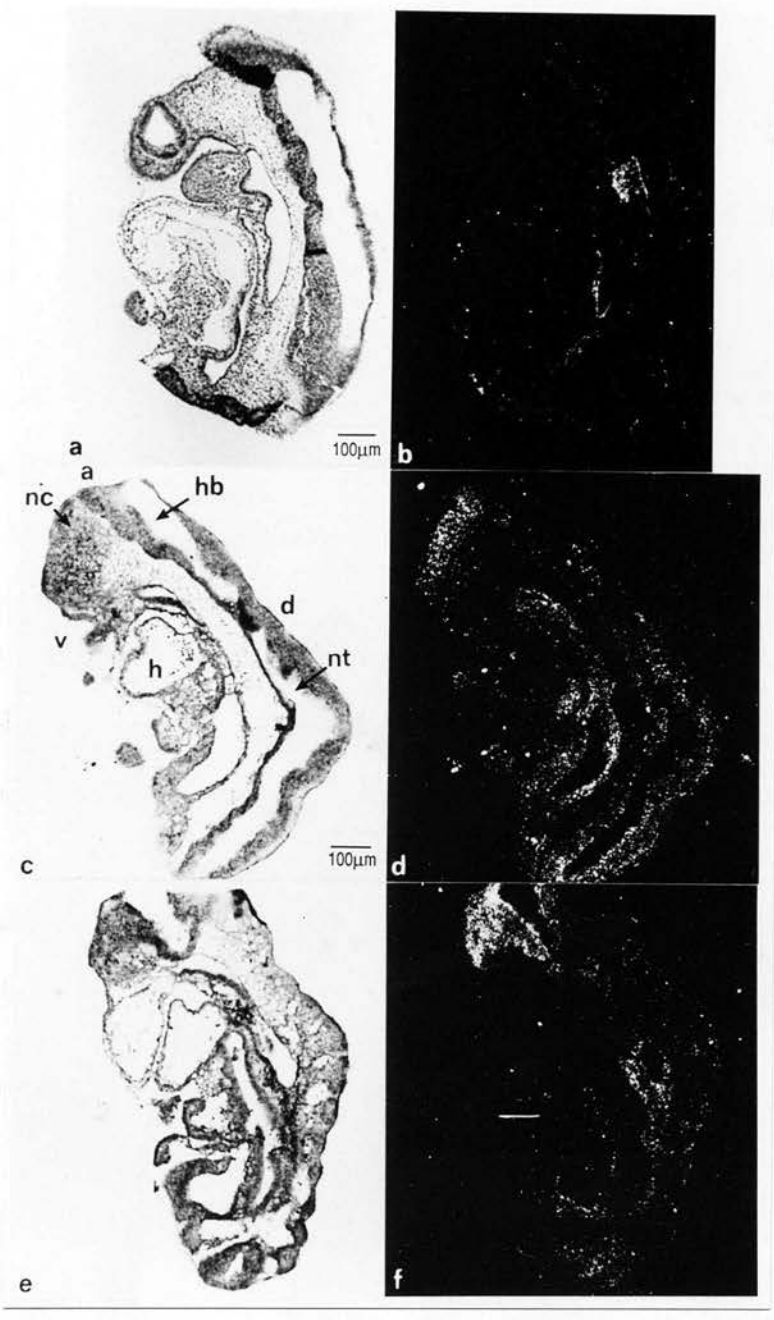


Figure 5.5. The rostral shift in position of the *Hox 2.9* expression domain in ot max (c-d) and ot 1 (e-f) embryos. (a-b) A control embryo section. (c-d) The ot max embryo section also shows the irregular undulations of the neural tube and hindbrain. a, anterior; nc, neural crest; hb, hindbrain; v, ventral; d, dorsal; h, heart; nt, neural tube.

5.2.3. Cells expressing Hox 2.9 *and* Krox 20 *are not as precisely localised within the hindbrain of RA treated embryos as in control embryos.*

The domain of *Hox 2.9* expression in ot 1 embryos lacks sharp expression boundaries (figure 5.3) and is therefore not as clearly defined as in the controls. During normal development *Hox 2.9* is not, at first (8 days), expressed in a sharply defined domain. Only later (~$8^3/_4$ days), when rhombomeres are clearly visible, does *Hox 2.9* expression become defined within sharp planar boundaries at the cellular level (chapter 4). It appears therefore that following RA treatment, eventhough *Hox 2.9* expression becomes localised within the hindbrain, the progression to a sharply defined expression pattern does not occur. In untreated embryos, both anterior and posterior boundaries of *Hox 2.9* expression coincide with equally well defined boundaries of *Krox 20* expression. The absence of cell mixing at these boundaries reflects the compartmentalisation of cells within the rhombomeres (Fraser *et al.*, 1990). In treated embryos only the posterior boundary of the *Hox 2.9* domain adjoins cells expressing *Krox 20*. Here, instead of the single planar boundary observed in controls there is an irregular alternation of cells expressing *Hox 2.9* and *Krox 20* resulting in a somewhat mosaic pattern at the interface (figure 5.6). As in control embryos, there does not appear to be intermingling of cells expressing these two genes so compartmentalisation of the cell types is maintained. However, regional localisation of the cells along the AP axis is not achieved following treatment.

At the anterior boundary of *Hox 2.9* expression in treated embryos, where the cells are not bounded by cells expressing *Krox 20*, the domain appears to extend more anteriorly (figure 5.3, 5.4). The expression here gradually decreases indicating perhaps a role for both *Hox 2.9* and *Krox 20* in defining their mutual expression limits. This could be achieved through repressive interraction between the genes or through cell movements resulting from repulsion between the two cell types. The latter is suggested by the situation at the posterior *Hox 2.9* boundary where cells are sorted out into regions of *Hox 2.9* and *Krox 20* expression in the absence of a well defined boundary. Analysis of hindbrain morphology (section 5.2.1) revealed that rhombomere like structures are present in the treated hindbrain. However, these are irregular and asymmetrical

and the absence of sharp planar expression boundaries in this region indicate that segmentation is not complete with the segments lacking complete definition.

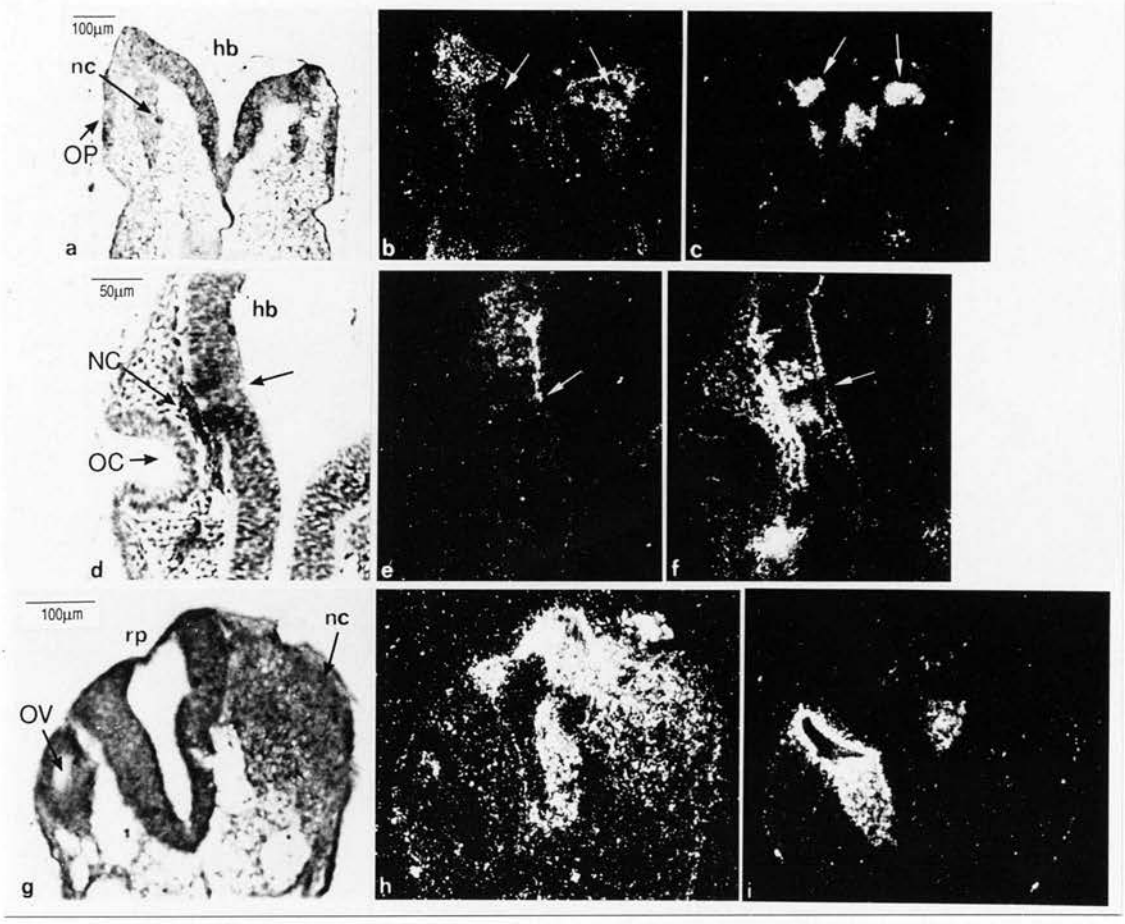


Figure 5.6. The expression domains of *Hox 2.9* and *Krox 20* in RA treated embryos lack sharp planar boundaries. (a-c) and (d-f) are adjacent frontal sections through two $9^1/_2$ day treated embryos. (g-i) Adjacent transverse sections through a $9^1/_2$ day treated embryo. b, e and h were probed with *Hox 2.9* and c, f and i were probed with *Krox 20*. hb, hindbrain; op, otic placode; nc, neural crest cells; oc, otic cup; rp, roof plate; ov, otic vesicle.

5.2.4. Expression of Hox 2.9 *and* Krox 20 *in neural crest cells.*

In untreated and treated embryos specific sets of neural crest cells that migrate from the hindbrain express *Hox 2.9* and *Krox 20* (figure 5.7). Only those which migrate from the region of *Hox 2.9* expression, express *Hox 2.9*. In controls the migrating neural crest cells just posterior to rhombomere 5 express *Krox 20* (Wilkinson *et al.*, 1989b). It is possible that these neural crest cells originate in rhombomere 5 but immediately take a posterior route. In treated embryos the neural crest cells that express *Krox 20* lie adjacent and posterior to the hindbrain domain of *Krox 20* expression, suggesting that they arise from the region expressing *Krox 20* in the hindbrain (figure 5.6). Labelling of neural crest cells with both genes is at a higher level and is more extensive in treated embryos than controls. However, the effect on *Krox 20* appears to be more extreme than on *Hox 2.9*. The elevated expression may be explained by a build up of neural crest cells in the mesoderm which is consistent with the previous observation that neural crest cell migration is retarded following RA treatment (Morriss & Thorogood, 1978).

5.2.5. Expression of Hox 2.9 in the trunk of the embryo is unaffected by excess RA.

The previous characterisation of the effects of excess RA treatment at the neural fold stage of rodent development (Morriss, 1972) reported that the trunk of the embryo and the post-cranial skeleton in the new born were relatively unaffected. *Hox 2.9* is normally expressed in a number of areas outside the hindbrain at $9^1/_2$ days including the posterior neural tube, the gut epithelium and the gut associated mesoderm. The expression of *Hox 2.9* in these regions seems to be unaffected by RA treatment (figure 5.4E).

5.2.6. Hox 2.9 expression in the hindbrain is prematurely down regulated following RA treatment.

Ot 1 embryos were also analysed at $10^1/_2$ days. At this stage *Krox 20* is no longer expressed in the hindbrain of untreated or treated embryos. *Hox 2.9* expression persists within rhombomere 4 of control embryos until 11 days. However, in treated embryos there is no

detectable expression of *Hox 2.9* at $10^1/_2$ days (figure 5.8). Thus *Hox 2.9* expression is prematurely down regulated following exposure to excess RA.

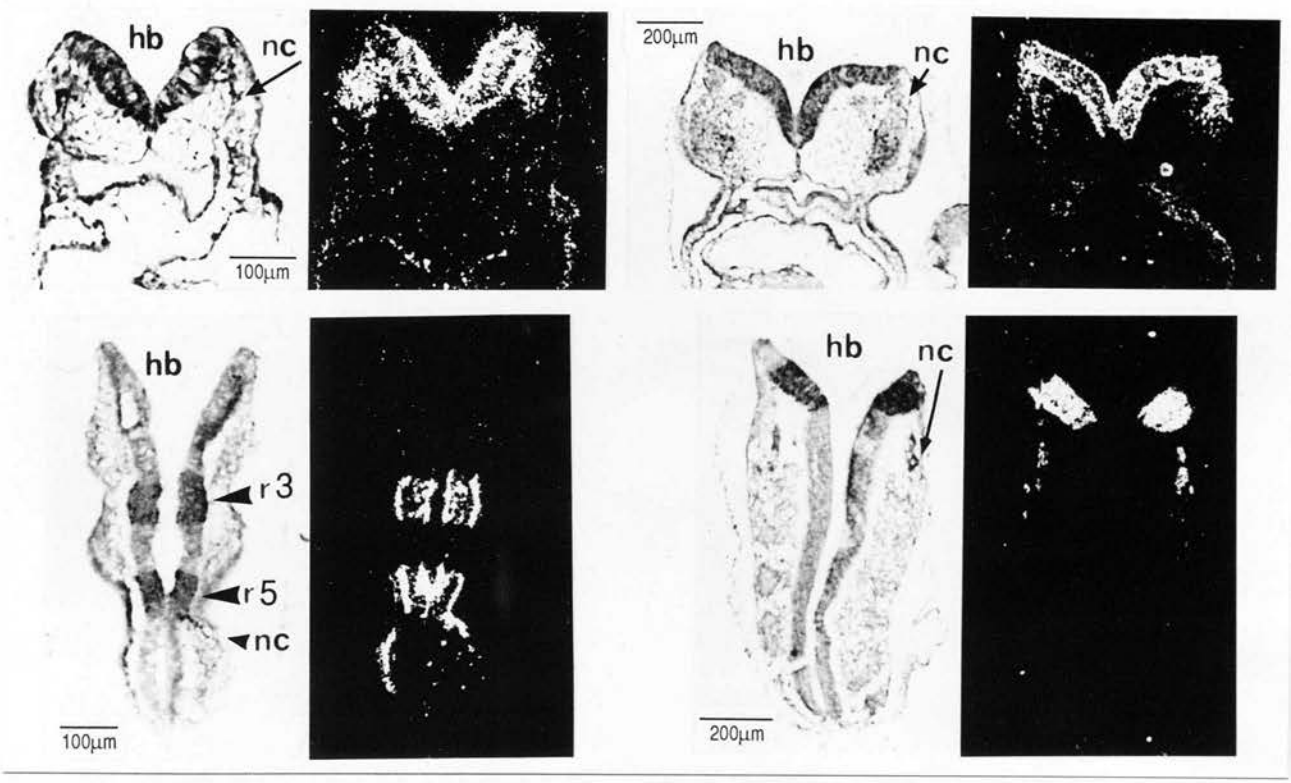


Figure 5.7. Expression of *Hox 2.9* and *Krox 20* in migrating neural crest cells in control and RA treated embryos. The top row of photo-micrographs shows expression of *Hox 2.9* in control (left) and treated (right) embryos. The bottom row shows expression of *Krox 20* in control (left) and treated (right) embryos. hb, hindbrain; nc, neural crest cells; r3, rhombomere 3; r5, rhombomere 5.

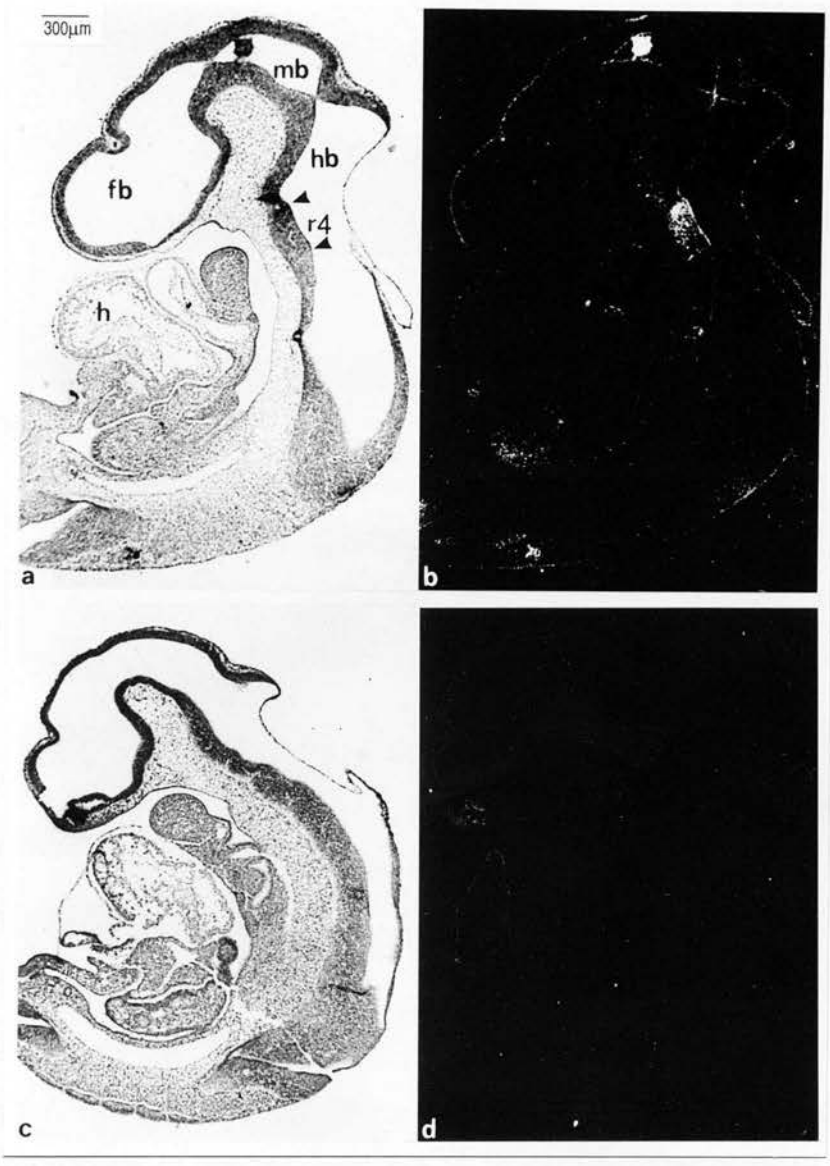


Figure 5.8. Expression of *Hox 2.9* in $10^1/_2$ day embryos treated with excess RA. (a-b) A sagittal section from a control embryo at $10^1/_2$ days. (c-d) A sagittal section from a RA treated embryo. fb, forebrain; mb, midbrain; hb, hindbrain; r4, rhombomere 4; h, heart.

5.2.7. Expression patterns are similarily disturbed in more severely affected embryos.

Ot max embryos, the more severely affected class, showed a similar pattern of expression of *Hox 2.9* and *Krox 20* to ot 1 embryos. In ot max embryos the midbrain, forebrain and cranial flexure are very much reduced (figure 5.2). The domains of expression of *Hox 2.9* and *Krox 20* occupy the same positions relative to the otic vesicle. However, with respect to the anterior tip of the brain they are shifted more anteriorly (figure 5.5). Therefore the forebrain and the midbrain are very much reduced. This implies that the more anterior a structure is the more severely it is affected by RA treatment. This may also explain why *Krox 20* is not expressed in the more anterior of its two normal expression domains.

5.3 Discussion.

5.3.1 The effect of excess RA on hindbrain segmentation increases our understanding of the normal process.

The organisation of the developing mouse hindbrain into segmental units along the AP axis and the segmental expression of *Hox 2.9* and *Krox 20* within the hindbrain have previously been described in detail (chapter 4). Here it is shown that the process of segmentation is disrupted following *in vivo* exposure to excess RA. In the normal embryo the segments are defined gradually in an anterior to posterior direction between 8 and 9 days of development. This is illustrated by the expression of *Hox 2.9* and *Krox 20,* which become localised to specific segments, with boundaries that become progressively sharper as the segments are defined (Murphy & Hill, 1991). Following treatment with RA during the initial stages of this process (8 days), the expression of *Hox 2.9* fails to become sharply defined, showing that the full definition of the segmental units is not achieved. The sharp planar boundary between the expression domains of *Hox 2.9* and *Krox 20*

in the normal embryo reflects the compartmentalisation of the hindbrain (Fraser *et al.*, 1990). Following RA treatment, cells expressing these two genes do not mix freely with each other so they remain compartmentalised to an extent. However, the lack of a planar boundary between the cells shows that their organisation into discrete domains along the AP axis is not achieved.

RA appears to directly affect the spatial organisation of the hindbrain. It has previously been suggested that the observed effects of RA on the anterior CNS of *Xenopus* may be mediated by the mesoderm as a result of a direct effect on gastrulation (Sive *et al.*, 1990). However in the experiments reported here the embryos were not treated until late in gastrulation. Similarily, Durston *et al.* (1989) showed that treatment of *Xenopus* embryos after gastrulation, continued to induce truncations at the anterior end of the AP axis. The treatment described here was given during segmentation of the hindbrain and is consistent with RA having a direct effect on the segmentation process. The ability of RA to interfere with the AP axis at the anterior end of mammalian and amphibian embryos, together with our knowledge of the behaviour of other systems in response to RA (section 1.2.6), indicates that RA may be an endogenous signalling substance in the cranial neural ectoderm. The results of exposure to excess RA can be seen as a distortion of this normal function.

The effect of RA on hindbrain segmentation may be telling us something about the normal segmentation process. It has been suggested that the cellular retinoic acid binding protein (CRABP) mediates the effect of RA in the embryo (Maden *et al.*, 1988). This is supported by the demonstration that the majority of RA in the embryo is CRABP bound (Dencker *et al.*, 1990). From the observation that the regions in which CRABP are expressed coincide with the regions most sensitive to RA treatment, Dolle *et al.* (1990; section 1.2.6) have suggested that CRABP expressing cells require a low level of RA, which is maintained by CRABP binding and inactivation. When excess RA is supplied, CRABP is incapable of maintaining the required low level due to saturation of the binding capacity. In addition they suggest that CRBP (cellular retinol binding protein) expressing cells produce locally required levels of RA from the metabolism of retinol. In the developing hindbrain both CRABP and CRBP are expressed in restricted domains

(Ruberte *et al.*, 1991). This may indicate that a low, tightly regulated, level of RA is required. The work of Dencker *et al.* (1990) demonstrated that CRABP protein is present at higher levels in rhombomeres 4 and 6 than in rhombomeres 1, 3 and 5 which indicates a role for fine differences in the level of binding protein, and therefore the balance of free RA, in defining segments. However, manipulation of the level of CRABP expression is required to determine if this is in fact the case.

As suggested in section 5.2.3 a careful examination of the distribution of cells expressing *Hox 2.9* and *Krox 20* in the treated embryo indicate that they may be mutually involved in defining their expression boundaries. Even though the cells are not correctly localised along the AP axis in the presence of excess RA, the cells expressing *Hox 2.9* and *Krox 20* do not freely mix. Furthermore the anterior boundary of *Hox 2.9* expression, where there is no expression of *Krox 20* in adjoining cells, is diffuse and imprecise indicating that repellent forces between the two cell types are involved in establishing the normal pattern.

5.3.2 The effect of excess RA on the expression of Hox 2.9 *and* Krox 20 *in the hindbrain.*

The normal expression patterns of *Hox 2.9* and *Krox 20* are disrupted following treatment with excess RA. It is not possible to say however if the effect on the genes is direct or indirect. It has also not been established whether the increase in *Hox 2.9* RNA levels following RA treatment of F9 cells (section 3.2.5) is a direct response. However, even though it is possible that *Hox 2.9* is a direct target of RA in F9 cells it may not be similarily responsive in the hindbrain. The interpretation of RA levels by cells in the hindbrain and by F9 cells may involve different receptor systems controlling different sets of responsive genes.

One interpretation of the current results is that RA disrupts the basic AP organisation of the hindbrain and as a result the information necessary to determine the precise expression of *Hox 2.9* and *Krox 20.* The level of RA and the expression of *Hox 2.9* and *Krox 20* may therefore be involved in the same multistep process of positional determination within the hindbrain.

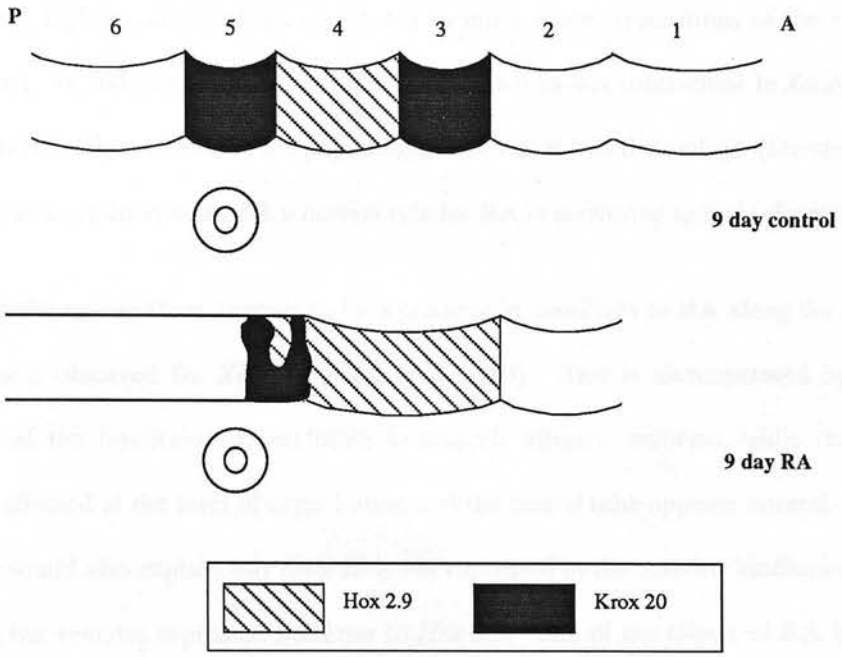


Figure 5.9 A diagramatic representation of the expression of *Hox 2.9* and *Krox 20* in the untreated (control) and RA treated 9 day, embryonic hindbrain. It shows the RA induced disruption of the segmental expression domains.

5.3.3 RA as a morphogen in the embryonic head.

The effect of RA on the AP axis of the developing embryo is dose dependent. This was shown here by the variability of effects on mouse embryos treated *in vivo* and also by exposure of *Xenopus* embryos to variable concentrations of RA *in vitro* (Sive *et al.*, 1990). In the *Xenopus* experiments, higher concentrations of RA led to more severe truncations of the embryos at the anterior end. In addition the minimum dose of RA that causes truncations in *Xenopus* was found to be approximately the same as the physiological level of RA in the embryo (Durston *et al.*, 1989). These features are consistent with a normal role for RA in conveying spatial information.

In the mouse there appears to be a gradient in sensitivity to RA along the AP axis of the CNS, as was observed for *Xenopus* (Sive *et al* 1990). This is demonstrated by the extreme reduction of the forebrain and midbrain in severely affected embryos, while the hindbrain is primarily affected at the level of organisation and the neural tube appears normal. A gradient in sensitivity would also explain why *Krox 20* is not expressed in the anterior hindbrain following RA treatment but remains expressed posterior to *Hox 2.9*. One of the effects of RA is the apparent posteriorisation of anterior structures where hindbrain features, including the expression of *Hox 2.9* and *Krox 20*, are found more anterior in treated embryos than in controls. In *Xenopus* the expression of the homeobox gene, *En*, is found nearer to the anterior end of the embryo following moderate levels of RA treatment (Sive *et al.*, 1990). A measure of the relative size of the various parts of the *Xenopus* CNS indicate that in treated embryos, while the overall size remains constant, the forebrain and midbrain are reduced and the hindbrain is enlarged (Durston *et al.*, 1989). A less direct experiment which supports this effect, is the treatment of midbrain cells in culture with RA. These begin to express a posterior specific homeobox gene *Hox 2.3* (Meijlink *et al.*, 1989). All of these results suggest that regions that would normally produce anterior structures become more posterior-like under the influence of excess RA.

The gradient in sensitivity to excess RA along the AP axis, and the transformation of anterior cells to a more posterior specification can be explained in terms of the disruption of a natural gradient, as was proposed by the effect of RA on the developing limb (Tickle *et al.*, 1982). As suggested by Dolle *et al.* (1990), CRABP and CRBP may be involved in determining a graded distribution of functional RA. However, we do not have enough information at present to draw definite conclusions about the nature of such a gradient. A gradient model was proposed by Durston *et al.* (1989) and extended by Sive *et al.* (1990) to explain the effects of RA on the amphibian AP axis. This model involves the induction of dorsal ectoderm by dorsal mesoderm to a graded series of potential anterior fates. RA is then a potential second factor which can subsequently determine a set of posterior lineages. The cells furthest away from the source of RA, or in some way experiencing lower levels of RA, would therefore maintain an anterior specification. The model explains the shift in the AP axis in terms of exposure of anterior cells to above normal levels of RA, the posterior specifier. A model involving high levels of RA in posterior cells fits with the recent observation that more anteriorly expressed genes in the human *HOX 2* cluster respond to lower levels of RA in teratocarcinoma cells (Simeone *et al.*, 1990). However, it has not been shown that the graded responsiveness of *HOX 2* cluster members plays a role in primary determination of the AP axis.

We find that only anterior structures are affected by excess RA in the mouse. Studies on *Xenopus* indicate that resistance to RA is acquired gradually in a posterior to anterior direction (Sive *et al.*, 1990). If this is also true in the mouse, we may be examining stages at which the AP axis in the trunk has already become resistant. Alternatively, RA may be involved in patterning the AP axis in the head alone. The neural crest cells provide evidence that specification in the head is different to specification in the trunk. Cranial neural crest cells appear to be prepatterned according to origin (Noden, 1983) whereas cells in the trunk are determined according to their destination (LeDourain, 1982, for review). The mesoderm appears to play a greater role in determining the AP axis in the trunk (Wedden *et al.*, 1988).

5.3.4 RA appears to primarily affect the dorsal ectoderm with neural crest cells mediating its effect on the mesoderm.

The early characterisation of the effects of RA on the developing mammalian embryo (Morriss, 1972; Morriss & Thorogood, 1978) indicated that the abnormalities observed in the cranial skeleton may be due to inappropriate migration of neural crest cells from the brain. Similarily the effect of RA on facial development of the chicken may be explained by a disturbance of neural crest cells (Wedden *et al.*, 1988). Only neural crest cells that arise from a single segmental unit of the mouse brain (rhombomere 4) express a specific DNA regulatory gene, *Hox 2.9* (chapter 4). As previously pointed out, this supports the idea of prepatterning of neural crest cells according to their origin, extending the segmental organisation of the head (Couly & LeDourain, 1990). *Krox 20* is also expressed in a specific set of neural crest cells although their precise origin has not been as easy to define (Wilkinson *et al.*, 1989a; section 5.2.4). Here we show that in RA treated embryos specific neural crest cells also express *Hox 2.9* and *Krox 20*. However the position of these cells is shifted anteriorly in relation to the pharyngeal arches, which is their destination in normal embryos (Tan & Morriss-Kay, 1986). Furthermore, there is an apparent build up of neural crest cells in the mesoderm adjacent to the hindbrain indicating that their migration is inhibited. It is therefore likely that these cells migrate to inappropriate facial primordia.

In general the results are consistent with a primary effect of excess RA exerted on the dorsal ectoderm with the effect on the mesenchyme being mediated by neural crest cells. However this is perhaps too simple. It has been reported that treatment of mammalian embryos with 13-cis-RA when the neural crest cells have already reached the facial primordia, leads to defects in the upper jaw and median cleft palate (Goulding & Pratt, 1986). This implies that RA may also directly affect the mesenchyme itself. From what we know about RA and how it is likely to be interpreted at the cellular level, it is no surprise that the effects appear complex, especially when we try to compare the results of different treatments at different stages of embryonic development in different systems. In this study the conclusions have been drawn from a consistent effect

following a standard treatment and indicate that the major effect of excess RA is a disturbance of the spatial organisation along the AP axis of the brain.

Chapter 6

Summary of conclusions and perspectives.

6.1 Conclusions.

The conclusions that can be drawn from the work presented in this thesis are outlined here. A more complete discussion of the work is found at the end of chapters 3, 4 and 5.

(1) A previously unidentified homeobox-containing cDNA clone, which maps to the Hox 2 cluster on chromosome 11, has been shown by sequence analysis to be most similar to *Hox 1.6*. The gene represented by this cDNA has been designated *Hox 2.9*.

(2) Compared with *Drosophila* homeobox-containing genes, *Hox 1.6* and *Hox 2.9* are most similar to a member of the *Antennapedia* complex, *labial*. This supports the hypothesis that vertebrate homeobox-containing gene clusters arose by duplication of a single ancestral cluster.

(3) *Hox 1.6* and *Hox 2.9* form the *labial*-like subfamily of homeobox-containing genes in the mouse. They are highly conserved within and outside the homeobox.

(4) *Hox 1.6* and *Hox 2.9* are expressed in a temporally and spatially restricted manner during development. Together with their similarity to *Drosophila* homeobox-containing genes, which are known to be developmental regulators, this indicates that they have roles in spatial determination during development.

(5) There are important similarities in the expression patterns of *Hox 2.9* and *Hox 1.6* indicating a functional similarity between the genes. With the striking exception of the hindbrain expression of *Hox 2.9* after $8^1/_2$ days, the expression patterns have the same AP restrictions indicating that they respond to at least some of the same signals in the embryo.

(6) *Hox 2.9* is expressed in rhombomere 4, a single segmental unit of the developing hindbrain, in the $8^1/_2$ to 11 day embryo. This indicates that the gene is involved in defining and/or identifying this segment during development and therefore offers a clear example of a mouse homeobox-containing gene being involved in the determination of a positional domain. This was the first demonstration of segmental expression of a vertebrate homeobox-containing gene.

(7) *Hox 2.9* is also uniquely expressed in the neural crest cells that migrate from rhombomere 4, indicating that neural crest cells are pre-patterned according to their rhombomeric origin in the hindbrain. This would extend the influence of hindbrain segmentation beyond the ectodermal germ layer.

(8) Looking more closely at the onset of segmental expression of *Hox 2.9* and the zinc finger encoding gene *Krox 20* in the developing hindbrain, it was shown that the domains are defined in an anterior to posterior direction and that segmental expression of both genes shortly preceeds the appearance of rhombomeres. The dynamics of the expression patterns therefore show that the genes are closely linked to the segmentation process and further indicate that the genes are involved in conferring positional information.

(9) Exposure of embryos to excess RA at approximately 8 days of development (when segments begin to be determined) disrupts the segmentation process. Clearly defined rhombomeres do not subsequently appear. This was shown morphologically and by the expression of *Hox 2.9* and *Krox 20* which lack sharp planar expression boundaries in treated embryos.

(10) Following RA treatment the expression domains of *Hox 2.9* and *Krox 20* are shifted anteriorly so that the *Hox 2.9* domain lies adjacent to the midbrain. *Krox 20* is expressed only in a single domain posterior to the *Hox 2.9* domain. The normal expression of *Krox 20* anterior to that of *Hox 2.9* is not detected.

(11) An analysis of the distribution of the different splicing products of the *Hox 1.6* gene revealed that the relative proportion of the two alternative transcripts changes as development proceeds. The longer transcript, which alone can produce a homeodomain-containing protein, appears to predominate at 8 days but makes up only a small proportion of the transcripts at 9 days.

6.2 Perspectives.

We are now approaching a very exciting time in the analysis of homeobox-containing genes and in the study of vertebrate development in general. The type of analysis presented in this thesis has provided important information about where and when these undoubtedly important genes are functional and therefore has indicated the particular embryonic systems in which they are involved. However, the next step is to determine what their precise roles are through observing the effect of gene manipulation. As described in section 1.4.4, some reports have been made on the effects of over-expression and ectopic expression of homeobox-containing genes (Wolgemouth *et al.*, 1989; Balling *et al.*, 1989) but these are difficult to interpret. Null mutations, where the endogenous genes are inactivated, are likely to be more informative with respect to the normal roles of the genes in the embryo. Although the inactivation of mammalian genes is a difficult task it now appears to be possible through the techniques of homologous recombination in cultured embryonic stem cells and the reintroduction of these cells into the embryo (see section 1.4.4). The procedures involved appear to be delicate and time consuming but they have now been successfully demonstrated for a number of genes (section 1.4.4) and the situation looks promising.

Hox 2.9 is a very interesting candidate for such analysis. It is involved in a now well defined segmentation process which can be observed morphologically, through the pattern of nerve growth and through segmental gene expression. Therefore, not only is an investigation of *Hox 2.9* function in the hindbrain of interest to determine how a homeobox gene is involved in the segmentation process, but the hindbrain should also prove to be particularly amenable to detailed analysis of the mutational effects. Problems may however arise from the fact that *Hox 2.9* is also expressed in other parts of the embryo and the absence of normal transcripts in these other sites may compound the mutant phenotype, making it difficult to specifically determine its role in the hindbrain. Also, since the gene is expressed in a large domain early in development, a null mutation may be lethal at early stages and the mutant embryos may be unrecoverable. It may take a number of carefully planned experiments, perhaps looking at the effect of mutations in other hindbrain genes in parallel, before the effects of the gene are clear.

Apart from creating mutations in segmentally expressed genes, other manipulations which alter the hindbrain segmentation pattern may yield information about the segmentation process and the role of segmentally expressed, potential genetic regulators, like *Hox 2.9*. Analysis of one such manipulation, i.e. exposure of embryos to excess RA, was reported in chapter 5 of this thesis. The altered expression patterns of the two genes analysed, *Hox 2.9* and *Krox 20*, revealed a number of features of the RA-treated hindbrain, including the fact that rhombomeres are not clearly defined and that there is a rostral shift in the position of expression domains within the hindbrain. The analysis should now be extended to observe the effect of excess RA on the expression of other hindbrain specific genes which may further our understanding of how RA disturbs segmentation. A number of interesting candidates for such analysis are described below.

(1) It was observed that the expression domain of *Hox 2.9* in treated embryos occupies a rhombomere-like structure adjacent to the hindbrain/midbrain junction. This is the normal position of rhombomere 1, which is three segments more anterior than the normal domain of *Hox 2.9*. In untreated embryos *En-2* is expressed at the hindbrain/midbrain junction including rhombomere 1 and the posterior midbrain (Davidson *et al.*, 1988). Therefore it would be worthwhile to determine if *En-2* expression overlaps *Hox 2.9* expression in treated embryos or if it is also shifted anteriorly.

(2) An interesting observation made in RA-treated embryos is that the otic vesicle and the *Hox 2.9* and *Krox 20* expression domains are co-ordinately shifted rostrally so that they approximately maintain their spatial relationship. The otic vesicle is thought to be initially induced by the underlying hindbrain and a role in this process has been proposed for the gene, *int-2*, (Wilkinson *et al.*, 1988). It is therefore important to look at the expression of *int-2* in treated embryos to determine if it is expressed adjacent to the shifted otic vesicle.

(3) A number of rhombomere boundaries are characterised by the expression pattern of homeobox-containing genes. The rhombomere 3 / rhombomere 2, rhombomere 5 / rhombomere 4 and rhombomere 6 / rhombomere 7 boundaries are the anterior expression boundaries of *Hox 2.8*, *Hox 2.7* and *Hox 2.6* respectively. Therefore an analysis of the expression patterns of these genes in the treated embryo would examine segmentation throughout the hindbrain.

The information required for correct spatial and temporal expression of homeobox-containing genes within the embryo is contained in adjacent control sequences (see section 1.4.3). It is therefore of interest to analyse and compare regions surrounding the genes to define these sequences. The control regions can be defined by assaying various portions of the flanking DNA for promoter activity. This can be done by attaching the region of interest to a convenient reporter gene, the expression of which is easily detected in transgenic mice. When the functional promoter regions have been defined these can be used to further manipulate gene expression.

Fusion-genes could be produced with the control regions of one gene and the coding sequences of another. These fusion genes would produce ectopic expression in transgenic mice, but in a more controlled way than in experiments previously described (Balling *et al.*, 1989). This type of experiment could be particularly interesting in examining the roles of *Hox 2.9* and *Krox 20* in the hindbrain. For example, a fusion-gene with the promoter region of *Hox 2.9* (or just the promoter elements that are responsible for rhombomere 4 expression, if they are definable and separable) attached to the coding sequence of *Krox 20* would spread the expression of *Krox 20* into rhombomere 4 where its effect on the endogenous expression of *Hox 2.9* and its effect on further development of the hindbrain could be observed. The converse experiment, with *Hox 2.9* driven by the *Krox 20* promoter, would also be of interest. The cloning of an 18.5kb *Hox 2.9* genomic clone with more than 10kb of DNA 5' of the coding sequence was reported in chapter 3. It may be possible to use this clone in such an experiment following definition of the promoter.

There are a number of genes that can be used as convenient reporter genes in assaying promoter activity. *lac Z* was used in the experiments described in section 1.4.3. Another alternative is the herpes simplex virus 1 thymadine kinase (HSV-1-TK) gene (Borrelli *et al.*, 1988; Heyman *et al.*, 1989), expression of which can be detected *in situ* by antibody staining. The use of HSV-1-TK in assaying *Hox 2.9* activity presents the possibility of another potentially interesting experiment, where the selective toxicity of the HSV-1-TK gene product could be used in attempting to ablate the cells of rhombomere 4. This would utilise the fact that HSV-1-TK is normally non-toxic to mammalian cells, but when nucleoside analogues, such as acyclovir, FIAU or

ganciclovir are available, they are metabolised to toxic intermediates (Borrelli *et al.*, 1988). This technique was successfully used to selectively kill transformed tissue culture cells (Borrelli *et al.*, 1988), to kill more than 99% of thymocytes in transgenic mice when the gene was fused to an immunoglobulin promoter (Heyman *et al.*, 1989) and to induce dwarfism when growth hormone producing cells were ablated (Borrelli *et al.*, 1989). In the experiment proposed here, pregnant germ-line transmitting transgenic mice carrying the HSV-1-TK gene on the *Hox 2.9* promoter (ideally, containing only the elements directing rhombomere 4 expression) would be treated with ganciclovir at the time that rhombomere 4 expression is clearly established (9 days) in order to specifically ablate rhombomere 4 cells. This experiment would be valuable in determining the effect of loss of a single complete segment on the developing brain and perhaps also in determining the specific fate of rhombomere 4. The removal of a single rhombomere from a living mouse embryo by microsurgery would be technically impossible.

There are however a number of potential problems that might interfere with the success of the experiment. Firstly, it has not been established if the nucleoside analogs can cross the placental barrier to reach the developing embryo and even if they do so, how long this process takes. Secondly, it may not be possible to define *Hox 2.9* elements that direct only rhombomere 4 expression and so it may be necessary to treat embryos which are also expressing the transgene in regions outside the brain. However, at 9 days the expression in the hindbrain is at a higher level than in the rest of the embryo and so it may be possible to see an independent effect on the hindbrain. Thirdly, in the event of successful ablation of rhombomere 4, the surrounding areas of the brain may simply compensate for the deleted region. However, this result would also be of interest.

It was observed in chapter 4 of this thesis that the relative abundance of *Hox 1.6* differential splicing products changes as development proceeds. However, it is not known what the relevance of this change is, since the function of the smaller, non-homeodomain encoding transcript has not been determined. It was speculated that this transcript may not have a function but may simply be the by-product of a post-transcriptional mechanism which silences the function

of the longer transcript. Alternatively, the shorter encoded protein may have a function, perhaps in interacting with the full-length homeodomain protein (see section 4.3.6). It is therefore important to determine if the shorter transcript is translated. Firstly it could be determined if the transcript can be translated in a cell-free system, but ultimately specific anyibodies would be required to determine if the protein is present in the embryo. This may not however be a simple task since problems have been encountered in raising antibodies specific to other homeobox-containing gene products (P. Budd, personal communication). In addition to discovering whether the transcript is translated, this antibody may also provide clues to the function of the protein since it would localise the gene product within the cell. Cellular localisation would determine if the protein is nuclear or cytoplasmic, indicating whether it is involved in DNA binding (which would be indirect since it does not have a DNA binding domain) or perhaps in interaction with other gene products in the cytoplasm.

In conclusion, the techniques of molecular biology have helped to greatly advance the field of developmental biology in recent years. In particular, they have facilitated the isolation and characterisation of a now large and growing number of developmentally important genes in higher organisms. *Hox 2.9* and *Hox 1.6* are interesting examples of such genes in the mouse. This thesis has described many aspects of their structure and expression. It has revealed that they are paralogous genes that have maintained many features in common but which have also specialised. The specialisation of *Hox 2.9* is shown very strikingly by its expression in the hindbrain. Further investigation of these genes is warranted to determine their specific roles in the developing embryo. Of particular interest is the role of *Hox 2.9* in the hindbrain. Advances are constantly being made that allow new and more searching experiments to be performed. These will tell us more about developmental genes and the ways in which they interact and function within the embryo.

Bibliography.

Acampora, D., D'Esposito, M., Faiella, A., Pannese, M., Migliaccio, E., Morelli, F., Stornaiuolo, A., Nigro, V., Simeone, A. & Boncinelli, E. (1989). The human *HOX* gene family. *Nucleic Acids Research* **17,** 10385-10402.

Adelmann, H.B. (1924). The development of the neural folds and cranial ganglia of the rat. *J. comp. Neurol.* **39,** 19-171.

Akam, M. (1987). The molecular basis of metameric pattern in the Drosophila embryo. *Development* **101,** 1-22.

Akam, M., Dawson, I. & Tear, G. (1988). Homeotic genes and the control of segment diversity. *Development* **104** Supplement, 123-133.

Akam, M. & Martinez-Arias, A. (1985). The distribution of *Ultrabithorax* transcripts in *Drosophila* embryos. *EMBO J.* **4,** 1689-1700.

Anderson, J.E., Ptashne, M. & Harrison, S.C. (1987). Structure of the repressor-operator complex of bacteriophage 434. *Nature* **326,** 846-852.

Anderson, K. (1987). Dorsal-ventral embryonic pattern genes of *Drosophila. Trends in Genetics* **3,** 91-97.

Anderson, W.F., Ohlendorf, D.H., Takeda, Y. & Matthews, B.W. (1981). Structure of the cro repressor from bacteriophage lambda and its interactions with DNA. *Nature* **290,** 754-758.

Asselineau, D., Bernard, B. A. Bailly, C. & Darmon, M. (1989). Retinoic acid improves epidermal morphogenesis. *Developmental Biology* **133,** 322-335.

Aviv, H. & Leder, P. (1972). Purification of biologically active globin messenger RNA by chromatography of oligothymidylic acid cellulose. *Proc. Natl. Acad. Sci U.S.A.* **69,** 1408-1412.

Awgulewitsch, A., Utset, M.F., Hart, C.P., McGinnis, W. & Ruddle, F.H. (1986). Spatial restriction in expression of a mouse homeobox locus within the central nervous system. *Nature* **320,** 328-335.

Baer, K.E.v. (1828). Uber die Entwicklungsgeschichte der Thiere, Bd. 1-3. Konigsberg.

Bailey, D.W. (1971). Recombinant inbred strains, an aid to finding identity, linkage and function of histocompatibility and other genes. *Transplantation* **11,** 325-327.

Baker, N.E. (1987b). Molecular cloning of sequences from *wingless*, a segment polarity gene in Drosophila: the spatial distribution of transcript in embryos. *EMBO J.* **6,** 1765-1773.

Balling, R., Deutsch, U. & Gruss, P. (1988). *undulated*, a mutation affecting the development of the mouse skeleton, has a point mutation in the paired box of *Pax-1. Cell* **55,** 531-535.

Balling, R., Mutter, G., Gruss, P. & Kessel, M. (1989). Craniofacial abnormalities induced by ectopic expression of the homeobox gene *Hox 1.1* in transgenic mice. *Cell* **58,** 337-347.

Baron, A., Featherstone, M.S., Hill, R.E., Hall, A., Galliot, B. & Duboule D. (1987). *Hox 1.6*: a mouse homeobox-containing gene member of the *Hox 1* complex. *EMBO J.* **6,** 2977-2986.

Bastian, H. & Gruss, P. (1990). A murine *even-skipped* homologue, *Evx-1*, is expressed during early embryogenesis and neurogenesis in a biphasic manner. *EMBO J.* **9,** 1839-1852.

Beeman, R.W. (1987). A homeotic gene cluster in the red flour beetle. *Nature* **327**, 247-249.

Beddington, R. (1982). An autoradiographic analysis of tissue potency in different regions of the embryonic ectoderm during gastrulation in the mouse. *J. Embryol. Exp. Morph.* **64**, 87-104.

Beddington, R. (1983). In *Development in Mammals*, Vol. 5, M.H. Johnson, editor. Elsevier Science Publishers B.V.

Beddington, R. (1987). In *Experimental Approaches to Mammalian Development*. Ed. by Pedersen, R.A and Rossant, J. Cambridge University Press, New York.

Benbrook, D., Lernhardt, E. & Pfahl, M. (1988). A new retinoic acid receptor identified from a hepatocellular carcinoma. *Nature* **333**, 669-672.

Bender, W., Akam, M., Karch, F., Beachy, P.A., Sperier, P., Lewis, E.B. & Hogness, D.S. (1983). Molecular genetics of the bithorax complex of *Drosophila melanogaster*. *Science* **221**, 23-29.

Benezra, R., Davis, R. L., Lockshon, D., Turner, D. L. & Weintraub, H. (1990). The protein Id: A negative regulator of helix-loop-helix DNA binding proteins. *Cell* **61**, 49-59.

Bodner, M., Castrillo, J-L., Theill, L.E., Deerinck, T., Ellisman, M. & Karin, M. (1988). The pituatory specific transcription factor GHF-1 is a homeobox-containing protein. *Cell* **55**, 505-518.

Bogarad, L.D., Utset, M.F., Awgulwitsch, A., Hart, C.P. & Ruddle, F.H. (1989). The developmental expression pattern of a new murine homeobox gene: *Hox 2.5*. *Dev. Biol.* **133**, 537-549.

Bolivar, F., Rodriguez, R.L., Greene, P.J., Betlach, M.C., Heyneker, H.l., Boyer, H.W., Crosa, J.H. & Falkow, S. (1977). Construction and characterisation of new cloning vehicles II. A multipurpose cloning system. *Gene* **2**, 95-113.

Boncinelli, E., Somma, R., Acampora, D., Pannese, M., D'Esposito, M. & Simeone, A. (1988). Organisation of human homeobox genes. *Hum. Reprod.* **3**, 880-886.

Bopp, D., Burri, M., Baumgartner, S., Frigerio, G. & Noll, M. (1986). Conservation of a large protein domain in the segmentation gene *paired* and in functionally related genes of *Drosophila*. *Cell* **47**, 1033-1040.

Borrelli, E., Heyman, R.A., Arias, C., Sawchenko, P.E. & Evans, R.M. (1989). Transgenic mice with inducible dwarfism. *Nature* **339**, 538-641.

Borrelli, E., Heyman, R., Hsi, M. & Evans, R.M. (1988). Targeting of an inducible toxic phenotype in animal cells. *Proc. Natl. Acad. Sci. USA* **85**, 7572-7576.

Boternbrood, E.C. & Nieuwkoop, P.D. (1973). The formation of the mesoderm in Urodelean amphibians. V. Its regional induction by the endoderm. *Wilhelm Roux' Arch. devl. Biol.* **173**, 319-332.

Brand, N., Petkovich, M., Krust, A., Chambon, P., de The, H., Marchio, A., Tillois,P. & Dejean, A. (1988). Identification of a second human retinoic acid receptor. *Nature* **332**, 850-853.

Breier, G., Dressler, G.R. & Gruss, P. (1988). Primary structure and developmental expression of *Hox 3.1*, a member of the murine *Hox 3* homeobox cluster. *EMBO J.* **7**. 1329-1336.

Brockes, J. P. (1989). Retinoids, homeobox genes, and limb morphogenesis. *Neuron* **2**, 1285-1294.

Brockes, J. P. (1990). Reading the retinoid signals. *Nature* **345**, 766-768.

Brown, R., Gray, R. H. & Bernstein, I. A. (1985). Retinoids alter the direction of differentiation in primary cultures of cutaneous keratinocytes. *Differentiation* **28**, 268-278.

Bucan, M., Yang-Feng, T., Colberg-Poley, A.M., Wolgemuth, D.J., Genet, J., Franke, U. & Lehrach, H. (1986). Genetic and cytogenetic localisation of the homeobox-containing genes on mouse chromosome 6 and human chromosome 7. *EMBO J.* **5**, 2899-2905.

Burglin, T.R., Finney, M., Coulson, A. & Ruvkun, G. (1989). *Caenorhabditis elegans* has scores of homeobox-containing genes. *Nature* **341**, 239-243.

Burmeister, M. & Lehrach, H. (1989). Isolation of large DNA fragments from agarose gels using agarase. *Trends Genet.* **5**, 41.

Carroll, S.B., Laymon, R.A., McCutcheon, M.A., Riley, P.D. & Scott, M.P. (1986). The localisation and regulation of *Antennapedia* protein expression in *Drosophila* embryos. *Cell* **47**, 113-122.

Carroll, S.B. & Scott, M.P. (1985). Localisation of *fushi tarazu* protein during *Drosophila* embryogenesis. *Cell* **43**, 47-57.

Carroll, S.B. & Scott, M.P. (1986). Zygotically active genes that affect the expression of the *fushi tarazu* segmentation gene during early *Drosophila* embryogenesis. *Cell* **45**, 113-126.

Cascio, S. & Gurdon, J.B. (1987). The initiation of new gene transcription during *Xenopus* gastrulation requires immediate preceding protein synthesis. *Development* **100**, 297--305.

Chavrier, P., Zerial, M., Lemaire, P., Almendral, J., Bravo, R. & Charnay, P. (1988). A gene encoding a protein with zinc fingers is activated during G0/G1 transition in cultured cells. *EMBO J.* **7**, 29-35.

Cho, K.W., Goetz, J., Wright, C.V., Fritz, A., Hardwicke, J. & DeRobertis, E.M. (1988). Differential utilisation of the same reading frame in a Xenopus homeobox gene encodes two related proteins sharing the same DNA-binding specificity. *EMBO J.* 7, 2139-2149.

Chytil, F. & Ong, D.E. (1984). Cellular retinoid binding proteins. In *The retinoids.* **vol 2** (eds M.B. Sporn, A.B. Roberts, D.S. Goodman) pp. 89-123. Academic, Orlando Florida.

Colberg-Poley, A.M., Voss, S.D., Chowdhury, K. & Gruss, P. (1985a). Structural analysis of murine genes containing homeobox sequences and their expression in embryonal carcinoma cells. *Nature* **314**, 713-718.

Colberg-Poley, A.M., Voss, S.D., Chowdhury, K., Stewart, C.L., Wagner, E.F. & Gruss, P. (1985b). Clustered homeoboxes are differentially expressed during murine development. *Cell* **43**, 39-45.

Coleman, K.G., Poole, S.J. Weir, M.P., Soeller, W.C. & Kornberg, T. (1987). The *invected* gene of *Drosophila*: sequence analysis and expression studies reveal a close kinship to the *engrailed* gene. *Genes & Development* **1**, 19-28.

Cooke, J. & Smith, J.C. (1989). Gastrulation and larval pattern in *Xenopus* after blastocoelic injection of a *Xenopus* inducing factor: experiments testing models for the normal organisation of mesoderm. *Devl Biol.* **131**, 383-400.

Cooke, J., Smith, J.C., Smith, E.J. & Yaqoob, M. (1987). The organisation of mesodermal pattern in *Xenopus laevis*: Experiments using a mesoderm inducing factor. *Development* **101**, 893-908.

Cote, S., Preiss, A., Haller, J., Schuh, R., Kiencin, A., Seifert, E. & Jackle, H. (1987). The *gooseberry* zipper region of *Drosophila*: five genes encode different spatially restricted transcripts in the embryo. *EMBO J.* **6**, 2793-2801.

Couly, G. & LeDouarin, N.M. (1990). Head morphogenesis in embryonic avian chimeras: evidence for a segmental pattern in the ectoderm corresponding to the neuromeres. *Development* **108**, 543-558.

Covell, D.A. & Noden, D.M. (1989). Embryonic development of the chick primary trigeminal sensory-motor complex. *J. Comp. Neurol.* **286**, 488-503.

Crick, F. (1969). Diffusion in embryogenesis *Nature* **225**, 420-422.

Dalcq, A. & Pasteels, J. (1937). Une conception nouvelle des bases physiologiques de la morphogenese. *Arch. Biol.* **48**, 669-710.

Dale, L. & Slack, J.M.W. (1987). Regional specification within the mesoderm of early embryos of *Xenopus laevis*. Development **100**, 279-295.

Dale, L., Matthews, G., Tabe, L. & Colman, A.C. (1989). Developmental expression of the protein product of *Vg1*, a localised maternal RNA in the frog *Xenopus laevis*. *EMBO J.* **8**, 1057-1065.

Davidson, D.R., Graham, E., Sime, C. & Hill, R.E. (1988). A gene with sequence similarity to *Drosophila engrailed* is expressed during the development of the neural tube and vertebrae in the mouse. *Development* **104**, 305-316.

Davis, C.A., Noble-Topham, S.E., Rossant, J. & Joyner, A.L. (1988). Expression of the homeobox containing gene En-2 delineates a specific region of the developing mouse brain. *Genes Dev.* **2**, 361-371.

Dearolf, C.R., Topol, J. & Parker, C.S. (1989). The *caudal* gene product is a direct activator of *fushi tarazu* transcription during Drosophila embryogenesis. *Nature* **341**, 340-343.

Dencker, L., Annerwall, E., Busch, C. & Eriksson, U. (1990). Localisation of specific retinoid-binding sites and expression of cellular retinoic-acid-binding protein (CRABP) in the early mouse embryo. *Development* **110**, 343-352.

Deol, M.S. (1964). The abnormalities of the inner ear in *kreisler* mice. *J. Embryol. exp. Morph.* **12**, 475-490.

DeRobertis, E. M., Oliver, G. & Wright, C. V. E. (1989). Determination of axial polarity in the vertebrate embryo: homeodomain proteins homeogenetic induction. *Cell* **57**, 189-191.

Desplan, C., Theis, J. & O'Farrell, P.H. (1985). The *Drosophila* developmental gene, *engrailed*, encodes a sequence specific DNA binding activity. *Nature* **318**, 630-635.

Deutsch, U., Dressler, G.R. & Gruss, P. (1988). *Pax 1*, a member of a paired box homologous murine gene family, is expressed in segmented structures during development. *Cell* **53**, 617-625.

Diederich, R.J. Merrill, V.K.L., Pultz, M.A. & Kaufman, T.C. (1989). Isolation, structure, and expression of *labial,* a homeotic gene of the *Antennapedia* complex involved in *Drosophila* head development. *Genes & Dev.* **3**, 399-414.

DiNardo, S., Kuner, J.M., Theis, J. & O'Farrell, P.H. (1985). Development of embryonic pattern in *Drosophila melanogaster* as revealed by accumulation of the nuclear *engrailed* protein. *Cell* **43**, 59-69.

Dolle, P., Izpisua-Belmonte, J-C., Falkenstein, H., Renucci, A. & Duboule, D. (1989a). Coordinate expression of the murine *Hox 5* complex homeobox-containing genes during limb pattern formation. *Nature* **342,** 767-772.

Dolle, P., Ruberte, E., Kastner, P., Petkovich, M., Stoner, C., Gudas, L.J. & Chambon, P. (1989b). Differential expression of genes encoding a, b and g retinoic acid receptors and CRABP in the developing limbs of the mouse. *Nature* **342**, 702-704.

Dolle, P., Ruberte, E., Leroy, P., Morriss-Kay, G. & Chambon, P. (1990). Retinoic acid receptors and cellular retinoid binding proteins. I. A systematic study of their differential pattern of transcription during mouse organogenesis. *Development* in press.

Dony, C. & Gruss, P. (1987). Specific expression of the *Hox 1.3* homeobox gene in murine embryonic structures originating from or induced by the mesoderm. *EMBO J.* **6**, 2965-2975.

Dressler, G.R., Deutsch, G.R., Chowdhury, K., Nornes, H.O. & Gruss, P. (1990).Pax-2, a new murine paired-box-containing gene and its expression in the developing excretory system. *Development* **109**, 787-795.

Driever, W. & Nusslein-Volhard, C. (1988). A gradient of *bicoid* protein in *Drosophila* embryos. *Cell* **54**, 83-93.

Driever, W. & Nusslein-Volhard, C. (1989). The *bicoid* protein is a positive regulator of *hunchback* transcription in the early *Drosophila* embryo. *Nature* **337**, 138-143.

Driever, W., Thoma, G. & Nusslein-Volhard, C. (1989). Determination of spatial domains of zygotic gene expression in the *Drosophila* embryo by the affinity of binding sites for the *bicoid* morphogen. *Nature* **340**, 363-367.

Duboule, D., Baron, A., Mahl, P. & Galliot, B. (1986). A new homeobox is present in overlapping cosmid clones which define the mouse Hox 1 locus. *EMBO J.* 5, 1973-1980.

Duboule, D. & Dolle, P. (1989). The structural and functional organisation of the murine *Hox* gene family resembles that of *Drosophila* homeotic genes. *EMBO J.* **5**, 1497-1505.

Duboule, D., Galliot, B., Baron, A. & Featherstone, M.S. (1989). Murine homeobox genes: some aspects of their organisation and structure. In *Cell to cell signals in mammalian development.* NATO ASI series, Vol. 26, pp. 97-108. S.W. deLaat, ed., Springer-Verlag, Berlin.

Duprey, P., Chowdhury, K., Dressler, G.R., Balling, R., Simon, D., Guenet, J-L. & Gruss, P. (1988). A mouse gene homologous to the *Drosophila* gene *caudal* is expressed in epithelial cells from the embryonic intestine. *Genes & Development* **2**, 1647-1645.

Durston, A. J., Timmermans, J. P. M., Hage, W. J., Hendricks, H. F. J., de Vries, N. J., Heidelveld, M. & Nieuwkoop, P. D. (1989). Retinoic acid causes an anteroposterior transformation in the developing central nervous system. *Nature* **340**,140-144.

Edgar, B.A., Weir, M.P., Schuibiger, G. & Kornberg, T. (1986). Repression and turnover pattern of *fusi tarazu* RNA in the early *Drosophila* embryo. *Cell* **47**, 747-754.

Evans, R.M. (1988). The steroid and thyroid hormone receptor superfamily. *Science* **240,** 889-895.

Fahrner, K., Hogan, B.L.M. & Flavell, R.A. (1987). Transcripts of *H-2* and *Qa* genes in embryonic and adult mice. *EMBO J* **6**, 1265-1271.

Fainsod, A., Bogarad, L.D., Ruusala, T., Crothers, D.M. & Ruddle, F.H. (1986). The homeo domain of a murine protein binds 5' to its own homeobox. *Proc. Natl. Acad. Sci. USA* **83**, 9532-9536.

Fainsod, A., Awgulewitsch, A. & Ruddle, F.H. (1987). Expression of the murine homeobox gene *Hox 1.5* during embryogenesis. *Dev. Biol.* **124**, 125-133.

Featherstone, M.S., Baron, A., Gaunt, S.J., Mattei, M-G. & Duboule, D. (1988). *Hox 5.1* defines a homeobox-containing gene locus on mouse chromosome 2. *Proc. Natl. Acad. Sci. USA* **85**, 4760-4764.

Feinberg, A. & Vogelstein, B. (1983). A technique for radiolabelling DNA restriction endonuclease fragments to high specific activity. *Anal Biochem.* **132**, 6-13.

Fibi, M., Zink, B., Kessel, M., Colberg-Poley, A.M., Labeit, S., Lehrach, H. & Gruss, P. (1988). Coding sequence and expression of the homeobox gene *Hox 1.3. Development* **102**, 349-359.

Fraser, S., Keynes, R. & Lumsden, A. (1990). Segmentation in the chick embryo hindbrain is defined by cell lineage restrictions. *Nature* **344**, 431-434.

Frigerio, G., Burri, M., Bopp, D., Baumgartner, S. & Noll, M. (1986). Structure of the segmentation gene *paired* and the *Drosophila* PRD gene set as part of a gene network. *Cell* **47**, 735-746.

Fritz, A.F., Cho, K.W.Y., Wright, C.V.E., Jegalian, B.G. & DeRobertis, E.M. (1989). Duplicated homeobox genes in *Xenopus*. *Dev. Biol.* **131**, 584-588.

Frohman, M. A., Boyle, M. & Martin, G. R. (1990). Isolation of the mouse *Hox 2.9* gene; analysis of embryonic expression suggests that positional information along the anterior-posterior axis is specified by mesoderm. *Development*110,, 589-607..

Frohnhofer, H.G. & Nusslein-Volhard, C. (1986). The organisation of the anterior pattern in the *Drosophila* embryo by the maternal gene *bicoid. Nature* **324**, 120-125.

Gallin, W.J., Chuong, C.M., Finkel, L.H. & Edelman, G.M. (1986). Antibodies to liver cell adhesion molecule perturb inductive interractions and alter fearher pattern and structure. *Proc, Natl. Acad Sci. USA* **83**, 8235-8239.

Garber, R.L., Kuroiwa, A. & Gehring, W. (1983). Genomic and cDNA clones of the homeotic locus of Antennapedia in *Drosophila. EMBO J.* **2**, 2027-2036.

Garcia-Bellido, A., Lawrence, P.A. & Morata, G. (1979). Compartments in animal development. *Scient. Am.* **241**, 102-110.

Gardner, R.L. (1982). Investigation of cell lineage and differentiation in the extraembryonic endoderm of the mouse embryo. *J. Embryol. Exp. Morphol.* **68**, 175-198.

Gaul, U., Seifert, E., Schuh, R. & Jackle, H. (1987). Analysis of *Kruppel* protein distribution during early *Drosophila* development reveals posttranscriptional regulation. *Cell* **50**, 639-647.

Gaunt, S.J. (1987). Homeobox gene *Hox 1.5* expression in mouse embryos: earliest detection by *in situ* hybridisation is during gastrulation. *Development* **101**, 51-60.

Gaunt, S. J. (1988). Mouse homeobox gene transcripts occupy different but overlapping domains in embryonic germ layers and organs: a comparison of *Hox 3.1* and *Hox 1.5*. *Development* **103**, 135-144.

Gaunt, S.J. Krumlauf, R. & Duboule, D. (1989). Mouse homeo-genes within a subfamily, *Hox 1.4, 2.6* and *5.1*, display similar anteroposterior domains of expression in the embryo, but show stage- and tissue-dependent differences in their regulation. *Development* 107, 131-141.

Gaunt, S.J., Miller, J.R., Powell, D.J. & Duboule, D. (1986). Homeobox gene expression in mouse embryos varies by the primitive streak stage. *Nature* **324**, 662-664.

Gaunt, S.J., Sharpe, P.T. & Duboule, D. (1988). Spatially restricted domains of homeo-gene transcripts in mouse embryos: relation to a segmented body plan. *Development* **104** supplement, 169-179.

Gaunt, S.J. & Singh, P.B. (1990). Homeogene expression patterns and chromosomal imprinting. *TIG* **6**, 208-212.

Gehring, W.J. (1985). Homeotic genes, the homeo-box and genetic control of development. *CSH Symp. Quant. Biol.* **50**, 243-251.

Gehring, W.J. (1987). Homeo boxes in the study of development. *Science* **236,** 1245-1252.

Gerhart, J. & Keller, R. (1986). Region specific cell activities in amphibian gastrulation. *Ann Rev. Cell Biol.* **2**, 201-229.

Gerhart, J., Vincent, J.P., Scharf, S.R., Black, S.D., Gimlich, R.L. & Danilchik, M. (1984). Localisation and induction in early development of *Xenopus. Phil. Trans. Roy. Soc. London Ser. B* **307**, 319-330.

Gerster, T., Balmaceda, C-G. & Roeder, R.G. (1990). The cell type-specific octamer transcription factor OTF-2 has two domains required for the activation of transcription. *EMBO J.* **9**, 1635-1643.

Gibson, G. & Gehring, W.J. (1988). Head and thoracic transformations caused by ectopic expression of *Antennapedia* during *Drosophila* development. *Development* **102**, 657-675.

Gibson, G., Schier, A., LeMotte, P. & Gehring, W.J. (1990). The specificities of *sex combs reduced* and *Antennapedia* are defined by a distinct portion of each protein that includes the homeodomain. *Cell* **62**, 1087-1103.

Giguere, V., Ong, E. S., Segui, P. & Evans, R. M. (1987). Identification of a receptor for the morphogen retinoic acid. *Nature* **330,** 624-629.

Godsave, S.F., Isaacs, H.V. & Slack, J.M.W. (1988). Mesoderm inducing factors: a small class of molecules. *Development* **102**, 555-566.

Goodrich, E.S. (1985). *Studies on the structure and development of vertebrates*, Dover publications.

Goulding, E. H. & Pratt, R. M. (1986). Isotretinonin teratogenicity in mouse whole embryo culture. *J. Craniofac. gen. Devl. Biol.* **6,** 99-112.

Goutte, C. & Johnson, A.D. (1988). a1 protein alters the DNA binding specificity of a2 repressor. *Cell* **52**, 875-882.

Graham, A., Papalopulu, N. & Krumlauf, R. (1989). The murine and *Drosophila* homeobox gene complexes have common features of organisation and expression. *Cell* **57**, 367-378.

Graham, A., Papalopulu, N., Lorimer, J., McVey, J.H., Tudenham, E.G.D. & Krumlauf, R. (1988). Characterisation of a murine homeobox gene, *Hox 2.6*, related to the *Drosophila Deformed* gene. *Genes Dev.* **2**, 1424-1438.

Green, M. (1981). *Genetic variants and strains of the laboratory mouse*. Fischer, Stuttgart.

Green, S. & Chambon, P. (1987). Oestradiol induction of a glucocorticoid-responsive gene by a chimaeric receptor. *Nature* **325**, 75-78.

Green, S. Chambon, P. (1988). Nuclear receptors enhance our understanding of transcription regulation. *TIG* **4**, 309-314.

Grunstein, M. & Hogness, D.S. (1975). Colony hybridisation: A method for the isolation of cloned DNAs that contain a specific gene. *Proc. Natl. Acad. Sci U.S.A.* **72**, 3961-3965.

Grunz, H. & Tacke, L. (1986). The inducing capacity of the presumptive endoderm of *Xenopus laevis* studied by trans-filter experiments. *Wilhelm Roux' Arch. devl Biol.* **192**, 130-137.

Gudas, L.J. & Wang, S.Y. (1986). *Prog. Clin. Biol. Res.* **226**, 181-189.

Gurdon, J.B. (1987). Embryonic induction - molecular prospects. *Development* **99**, 285-306.

Gurdon, J.B. (1988). A community effect in animal development. *Nature* **336**, 772-774.

Gurdon, J.B. (1989). The localisation of an inductive response. *Development* **105**, 27-33.

Hafen, E., Kuroiwa, A. & Gehring, W.J. (1984a). Spatial distribution of transcripts from the segmentation gene *fushi tarazu* during *Drosophila* embryonic development. *Cell* **37**, 833-841.

Hafen, E., Levine, M. & Gehring, W.J. (1984b). Regulation of *Antennapedia* transcript distribution by the *bithorax* complex in *Drosophila. Nature* **307**, 287-289.

Hamburger, V. (1988). *The heritage of experimental embryology. Hans Spemann and the organiser.* Oxford University Press.

Han, K., Levine, M.S. & Manley, J.L. (1989). Synergistic activation and repression of transcription by *Drosophila* homeobox proteins. *Cell* **56**, 573-583.

Hanahan, D. (1985). Techniques for transformation of *E.coli*. In *DNA cloning a practical approach* Vol 1 (ed. D.M. Glover) IRL Press, Oxford **1**, 109.

Haneman, E., Trevarrow, B., Metcalfe, W.K., Kimmel, C.B. & Westerfield, M. (1988). Segmental pattern of development of the hindbrain and spinal cord of the zebrafish embryo. *Development* **103**, 49-58.

Hanes, S.D. & Brent, R. (1989). DNA specificity of the *Bicoid* activator protein is determined by homeodomain recognition helix residue 9. *Cell* **57**, 1275-1283.

Harding, K., Wedeen, C., McGinnis, W. & Levine, M. (1985). Spatially regulatede expression of homeotic genes in *Drosophila. Science* **229**, 1236-1242.

Hart, C.P., Awgulewitsch, A., Fainsod, A. & Ruddle, F.H. (1985). 'Homeobox gene complex on mouse chromosome 11: molecular clonong, expression in embryogenesis, and homology to a human homeobox locus.' *Cell* **43**, 9-18.

Hart, C.P., Fainsod, A. & Ruddle, F.H. (1987). Sequence analysis of the murine *Hox 2.2, 2.3* and *2.4* homeoboxes: evolutionary and structural comparisons. *Genomics* **1**, 182-185.

Hauser, C.A., Joyner, A.L., Klein, R.D., Learned, T.K., Martin, G.R. & Tjian, R. (1985). Expression of homologous homeobox-containing genes in differentiated teratocarcinoma cells and mouse embryos. *Cell* **43**, 19-28.

Hertwig, P. (1942). Neue Mutationen und Koppelungsgruppen bei der Hausmaus. *Z. Indukt. Abstammungs-Vererbungsl.* **80**, 220-246.

Heyman, R.A., Borrelli, E., Lesley, J., Anderson, D., Richman, D.D., Baipd, S.M., Hyman, R. & Evans, R.M. (1989). Thymidine kinase obliteration: creation of transgenic mice with controlled immune deficiency. *Proc. Natl. Acad. Sci. USA* **86**, 2698-2702.

Hill, R.E., Hall,C.M., Sime, C.M. & Hastie, N.D. (1987). A mouse homeobox-containing gene maps near to a developmental mutation. *Cytogenet. Cell. Genet.* 44, 171-174.

Hill, R.E., Jones, P.F., Rees, A.R., Sime, C.M., Justice, M.J., Copeland, N.G., Jenkins, N.A., Graham, E. & Davidson, D.R. (1989). A new family of mouse homeobox-containing genes: molecular structure, chromosomal location, and developmental expression of *Hox 7. Genes & Development:* **3,** 26-37.

Hill, R.E., Shaw, P.H., Barth, R.K. & Hastie,N.D. (1985). A genetic locus closely linked to a protease inhibitor gene complex controls the level of multiple mRNA transcripts. *Mol. Cell. Biol.* **5,** 2114-2122.

Hoey, T. & Levine, M. (1988). Divergent homeo-box proteins recognise similar DNA sequences in Drosophila. *Nature* **332**, 858-861.

Hogan, B.H., Holland, P. & Schofield, P. (1985). How is the mouse segmented? *TIG* **1**, 67-74.

Hogan, B.H., Costantini, F. & Lacy, E. (1986) *Manipulating the mouse Embryo, A Laboratory Manual*. Cold Spring Harbour Laboratory Press, Cold Spring Harbour, New York.

Holland, P.W.H. & Hogan, B.L.M. (1988a). Spatially restricted patterns of expression of the homeobox containing gene *Hox 2.1* during mouse embryogenesis. *Development* **102,** 159-174.

Holland, P.W.H. & Hogan, B.L.M. (1988b). Expression of homeobox genes during mouse development: a review. *Genes & Development.* **2,** 773-782.

Hooper, M., Hardy, K., Handyside, A., Hunter, S. & Monk, M. (1987). HPRT deficient Lesch-Nyhan mouse embryos derived from germline colonisation by cultured cells. *Nature* **326**, 292-295.

Hopwood, N.D., Pluck, A. & Gurdon, J.B. (1989). A *Xenopus* mRNA relayed to *Drosophila twist* is expressed in response to induction in the mesoderm and the neural crest. *Cell* **59,** 893-903.

Howard, K. & Ingham, P. (1986). Regulatory interractions between the segmentation genes *fushi tarazu, hairy* and *engrailed* in the *Drosophila* blastoderm. *Cell* 44, 949-957.

Ingham, P. (1985). The regulation of the Bithorax complex. *Trends in Genetics* **1**, 112-116.

Ingham, P. (1988). The molecular genetics of embryonic pattern formation in *Drosophila*. *Nature* **335**, 25-34.

Jackle, H., Tautz, D., Schuh, R., Seifert, E. & Lehmann, R. (1986). Cross regulatory interractions among the gap genes of *Drosophila. Nature* **324**, 668-670.

Jackson, I. (1989). In *Genes and Embryos*. Ed Glover, D.M. & Hames, B.D. IRL Press at Oxford University Press.

Jackson, I.J., Schofield, P. & Hogan, B. (1985). A mouse homeobox gene is expressed during embryogenesis and in adult kidney. *Nature* **317**, 745-748.

Jacobson, A. (1966). Inductive processes in embryonic development. *Science* **152**, 25-34.

Jaynes, J.B. & O'Farrell, P.H. (1988). Activation and repression of transcription by homeodomain-containing proteins that bind a common site. *Nature* **336**,744-749.

Johnson, A.D. & Herskowitz, I. (1985). A repressor (Mat a2 product) and its operator control expression of a set of cell type specific genes in yeast. *Cell* **42**, 237-247.

Johnson, M.H. & Ziomek, C.A. (1981). The formation of two distinct cell lineages within the mouse morula. *Cell* **24**, 71-80.

Joyner, A.L. & Martin, G.M. (1987). *En-1* and *En-2*, two mouse genes with sequence homology to the *Drosophila engrailed* gene: expression during embryogenesis. *Genes & Development* **1**, 29-38.

Joyner, A.L., Skarnes, W.C. & Rossant, J. (1989). Production of a mutation in the mouse *En-2* gene by homologous recombination in embryonic stem cells. *Nature* 338, 153-156.

Jurgens, G., Wieschaus, E., Nusslein-Volhard, C. & Kluding, H. (1984). Mutations affecting the pattern of the larval cuticle in *Drosophila melanogaster.* II Zygotic loci on the third chromosome. *Wilhelm Roux Arch. devl. Biol.* **193**, 283-295.

Kalln, B. (1953). On the significance of the neuromeres and similar structures in vertebrate embryos. *J. Embryol. exp. Morph.* **1**, 387-392.

Kao,, K.R., Masui, Y. & Elinson, R.P. (1986) Lithium induced respecification of pattern in *Xenopus laevis* embryos. *Nature* **322**, 371-373.

Karkinen-Jaaskelainen, M. (1978). Transfilter lens induction in avian embryo. *Differentiation* **12,** 31-37.

Kastner, P., Krust, A., Mendelsohn, C., Garnier, J-M., Zelent, A., Leroy, P., Staub, A. & Chambon, P. (1990). Murine isoforms of retinoic acid receptor-g with specific patterns of expression. *Proc. Natl. Acad. Sci. USA* **87**, 2700-2704.

Kaufman, T.C., Lewis, R. & Wakimoto, B. (1980). Cytogenetic analysis of chromosome 3 in *Drosophila melanogaster*: the homeotic gene complex in polythene chromosome interval 84A-B. *Genetics* **94,** 115-133.

Kelleher, C.A., Goutte, C. & Johnson, A.D. (1988). The yeast cell type specific repressor a2 acts co-operatively with a non cell type specific protein. *Cell* **53**, 927-936.

Kelly, S.J. (1977). Studies on the developmental potential of 4- and 8-cell stage mouse embryos. *J. Exp. Zool.* **200**, 365-376.

Kessel, M., Balling, R. & Gruss, P. (1990). Variations of cervical vertebrae after expression of *Hox 1.1* transgene in mice. *Cell* **61**, 301-308.

Kessel, M. & Gruss, P. (1990). Murine developmental control genes. *Science* **249**, 374-379.

Kessel, M., Schulze, F., Fibi, M. & Gruss, P. (1987). Primary structure and nuclear localisation of a murine homeodomain protein. *Proc. Natl. Acad. Sci. USA* **84**, 5306-5310.

Keynes, R.J. & Stern, C.D. (1984). Segmentation in the vertebrate nervous system. *Nature* **310**, 786-789.

Kieny, M., Mauger, A. & Sengel, P. (1972). Early regionalisation of the somitic mesoderm as studied by the development of the axial skeleton of the chick embryo. *Dev Biol.* **28**, 142-161.

Kilcherr, F., Baumgartner, S., Bopp, D., Frei, E. & Noll, M. (1986). Isolation of the *paired* gene of *Drosophila* and its spatial expression during early embryogenesis. *Nature* **321**, 493-499.

Kimelman, D., Abraham, J.A., Haaparanta, T., Palisi, T.M. & Kirschner, M. (1988). The presence of FGF in the frog egg: its role as natural mesoderm inducer. *Science* **242**, 1053-1056.

Kimelman, D. & Kirschner, M. (1987). Synergistic induction of mesoderm by FGF and TGFb and the identification of an mRNA coding for FGF in the early *Xenopus* embryo. *Cell* **51**, 369-377.

Kimmel, C.B., Sepich, D.S. & Trevarrow, B. (1988). Development of segmentation in zebrafish. *Development* **104** (supplement), 197-207.

Ko, H.S., Fast, P., McBride, W. & Staudt, L.M. (1988). A human protein specific for the immunoglobulin octamer DNA motif contains a functional homeobox domain. *Cell* **55**, 135-144.

Kochar, D. M. (1973). Limb development in mouse embryos. I. Analysis of teratogenic effects of retinoic acid. *Teratology* **7**, 289-298.

Koller, B.H., Hagemann, L.J., Doetschman, T., Hagaman, J.R., Huang, S., Williams, P.J., First, N.C., Maeda, N. & Smithies, O. (1989). Germ line transmission of a planned alteration made in a hypoxanthine phosphoribosyltransferase gene by homologous recombination in embryonic stem cells. *Proc. Natl. Acad. Sci. USA* **86**, 8927-8931.

Krasnow, M.A., Saffman, E.E., Kornfeld, K. & Hogness, D.S. (1989). Transcriptional activation and repression by *Ultrabithorax* proteins in cultured *Drosophila* cells. *Cell* **57**, 1031-1043.

Krieg, P.A. & Melton, D.A. (1984). Formation of the 3' end of histone mRNA by post-transcriptional processing. *Nature* **308**, 203-206.

Krumlauf, R., Holland, P.W.H., McvVey, J.H. & Hogan, B.L.M. (1987). Developmental and spatial patterns of expression of the mouse homeobox gene, *Hox 2.1*. *Development* **99**, 603-617.

Krust, A., Kastner, Ph., Petkovich, M., Zelent, A. & Chambon, P. (1989). A third human retinoic acid receptor, hRARg. *Proc. Natl. Acad. Sci. USA* **86**, 5310-5314.

Kuehn, M.R., Bradley, A., Robertson, E.J. & Evans, M.J. (1987). A potential animal model for Lesch-Nyam syndrome through introduction of HPRT mutations into mouse. *Nature* **326**, 295-298.

Kuratani, S., Tanaka, S., Ishikawa, Y. & Zukeran, C. (1988). Early development of the facial nerve in the chick embryo with special reference to the development of the *chorda tympani*. *Am. J. Anat.* **182**, 169-182.

Kurihara, K. & Sasaki, N. (1981). Transmission of homeogenetic induction in presumptive ectoderm and newt embryo. *Dev. Growth Diff.* **23**, 361-369.

Kuriowa, A., Hafen, E., Gehring, W.G. (1984). Cloning and transcriptional analysis of the segmentation gene *fushi tarazu* of *Drosophila*. *Cell* **37**, 825-831.

Kuziora, M.A. & McGinnis, W. (1988). Autoregulation of a *Drosophila* selector gene. *Cell* **55**, 477-485.

Kuziora, M.A. & McGinnis, W. (1989). A homeodomain substitution changes the regulatory specificity of the *Deformed* protein in *Drosophila* embryos. *Cell* **59**, 563-571.

Lammer, E. J., Chen, D. T., Hoar, R. M., Agnish, N. D., Benke, P. J., Braun, J. T., Curry, C. J., Fernhoff, P. M. Grix, A. W., Lott, I. T., Richard, J. M. & Sun, S. C. (1985). Retinoic acid embrtopathy. *The New England J. of Medicine* **313,** 837-841.

LaRosa, G.L. & Gudas, L.J. (1988a). An early effect of retinoic acid: Cloning of an mRNA (Era-1) exhibiting rapid and protein synthesis-independent induction during teratocarcinoma stem cell differentiation. *Proc. Natl. Acad. Sci. USA* **85**, 329-333.

LaRosa, G.J. & Gudas, L.J. (1988b). Early retinoic acid induced F9 teratocarcinoma stem cell gene ERA-1: alternate splicing creates transcripts for a homeobox-containing protein and one lacking the homeobox. *Mol. & Cell Biol.* **8,** 3906-3917.

Laughon, A. & Scott, M.P. (1984). Sequence of a *Drosophila* segmentation gene: protein structure homology with DNA binding proteins. *Nature* **310**, 25-31.

Lawrence, P.A. (1987). Pair-rule genes: Do they paint stripes or draw lines? *Cell* **51,** 879-800.

Le Douarin, N. (1964). Induction de l'endoderme pre-hepatique par le mesoderme de l'aire cardiaque chez l'embryon de poulet. *J. Embryol. exp. Morph*. **12**, 651-664.

Le Douarin, N. M. (1982). *The neural crest.* pp259. Cambridge: Cambridge university press.

Leivo, I., Vaheri, A., Tompl, R. & Wartiovaara, J. (1980). Appearance and distribution of collagens and laminin in the early mouse embryo. *Dev Biol.* **76**, 100-114.

Le Mouellic, H., Condamine, H. & Brulet, P. (1988). Pattern of transcription of the homeo gene *Hox 3.1* in the mouse embryo. *Genes Dev.* **2**, 125-135.

Lewis, E. B. (1978). A gene complex controlling segmentation in *Drosophila*. *Nature* **276,** 565-570.

Lewis, J. H., Slack, J.M.W. & Wolpert, L. (1977). Thresholds in development. *J. Theor. Biol.* **65,** 579-590..

Ling, N., Ying, S-Y., Ueno, N., Shimasaki, S., Esch, F., Hotta, M. & Guillemin, R. (1986). Pituatary FSH is released by a heterodimer of the ß-subunits from the two forms of inhibin. *Nature* **321**, 779-782.

Lotan, R. (1980). Effects of vitamin A and its analogues (retinoids) on normal and neoplastic cells. *Biochim. Biophys. Acta* **605,** 33-91.

Lumsden, A. & Keynes, R. (1989). Segmental patterns of neuronal development in the chick hindbrain. *Nature* **337**, 424-428.

Lyon, M.F. & Searle, A.G. (1989). *Genetic variants and strains of the Laboratory mouse*. Oxford University Press, Oxford.

Macdonald, P.M., Ingham, P. & Struhl, G. (1986). Isolation structure and expression of *even-skipped:* a second pair-rule gene of *Drosophila* containing a homeobox. *Cell* **47**, 721-734.

Macdonald, P.M. & Struhl, G. (1986). A molecular gradient in early *Drosophila* embryos and its role in specifying the body pattern. *Nature* **324**, 537-545.

Maden, M. (1982). Vitamin A and pattern formation in the regenerating limb. *Nature* **295**, 672-675.

Maden, M. (1985). Retinoids and the control of pattern in limd development and regeneration. *TIG* **1**, 103-107.

Maden, M., Ong, D.E., Summerbell, D. & Chytil, F. (1988). Spatial distribution of a cellular protein binding to retinoic acid in the chick limb bud. *Nature* **335**, 733-735.

Maden, M., Ong, D.E., Summerbell, D., Chytil, F. & Hirst, E.A. (1989). Cellular retinoic acid-binding protein and the role of retinoic acid in the development of the chick embryo. *Devl Biol.* **135**, 124-132.

Maestro, A.M., Babich, M.A., Taylor, W.D. & Keith, A.D. (1984). Diffusion of a small molecule in the cytoplasm of mammalian cells. *Proc. Natl. Acad. Sci. U.S.A.* **81**, 3414-3418.

Mahon, K.A., Westphal, H. & Gruss, P. (1988). Expression of homeobox gene *Hox 1.1* during mouse embryogenesis. *Development* **104** supplement, 187-195.

Manglesdorf, D. J., Ong, E. S., Dyck, J. A. & Evans, R. M. (1990). Nuclear receptor that identifies a novel retinoic acid response pathway. *Nature* **345**, 224-229.

Mann, R.S. & Hogness, D.S. (1990). Functional dissection of *Ultrabithorax* proteins in *D. melanogaster. Cell* **60**, 597-610.

Martinez-Arias, A. (1986). The *Antennapedia* gene is required and expressed in parasegments 4 and 5 of the *Drosophila* embryo. *EMBO J.* **5**, 135-141.

Martinez-Arias, A. & Lawrence, P.A. (1985). Parasegments and compartments in the *Drosophila* embryo. *Nature* **313**, 639-642.

Mason, A.J., Berkenmeier, L., Schmelzer, C. & Schwall, R. (1989). Activin B: precursor sequences, genomic structure and in vitro activities. *Mol. Endocrinol.* **3**, 1352-1358.

Matthews, B.W., Ohlendorf, D.H., Takeda, Y. & Anderson, W.F. (1982). Structure of the DNA binding region of lac repressor inferred from its homology with cro repressor. *Proc. Natl. Acad. Sci USA.* **79**, 1428-1432.

Mavilio, F., Simeone, A., Boncinelli, E. & Andrews, P.W. (1988). Activation of four homeobox gene clusters in human embryonal carcinoma cells induced to differentiate by retinoic acid. *Differentiation* **37**, 73-79.

Maynard Smith, J. (1960). Continuous, quantized and normal variation. *Proc. R. Soc. B.* **152**, 397-409.

McGinnis, W. (1985). Homeo-box sequences of the Antennapedia class are conserved only in higher animal genomes. *CSH Symp. Quant. Biol.* **50**, 263-270.

McGinnis, W., Levine, M.S., Hafen, E., Kuroiwa, A. & Gehring, W.J. (1984). A conserved DNA sequence in homoeotic genes of the *Drosophila Antennapedia* and *bithorax* complexes. *Nature* **308**, 428-433.

McKay, R.D.G. (1989). The origins of cellular diversity in the mammalian central nervous system. *Cell* **58**, 815-821.

McKay, D.B. & Steitz, T.A. (1981). Structure of catabolite gene activator protein at 2.9A resolution suggests binding to left handed B-DNA. *Nature* **290**, 744-749.

McLaren, A. (1976). *Mammalian Chimaeras*. Cambridge University Press, Cambridge.

McMahon, A.P. & Bradley, A. (1990). The *Wnt*-1 (*int*-1) proto-oncogene is required for development of a large region of the mouse brain. *Cell* **62**, 1073-1085.

Mead, D.A., Szczesna-Skorupa, E. & Kemper, B. (1986). Single stranded DNA 'blue' T7 promoter plasmids: a versatile tandem promoter system for cloning and protein engineering. *Protein Engineering* **1**, 67-74.

Meier, S. & Tam, P.P.L. (1982). Metameric pattern development in the embryonic axis of the mouse. I. Differentiation of the cranial segments. *Differentiation* **21**, 109-122.

Meijlink, F., de Laaf, L., Verrikzer, P., de Graaf, W. & Deschamps, J. (1989) Regulation of expression of the *Hox 2.3* gene. In *Cell to cell signals in mammalian development.* (ed. S. W. de Laat), vol. H26, pp. 23-41. NATO ASI Series, Springer-Verlag, Berlin/Heidelberg.

Melton, D.A., Kreig, P.A., Rebagliati, M.R., Maniatis, T., Zinn, K. & Green, M.R. (1984). Effective *in vitro* synthesis of biologically active RNA and RNA hybridisation probes from plasmids containing a bacteriophage SP6 promoter. *Nucl. Acids. Res.* **12**, 7035-7056.

Merrill, V.K., Diederich, R.J., Turner, F.R. & Kaufman, T.C. (1989). A genetic and developmental analysis of mutations in *labial*, a gene necessary for proper head formation in *Drosophila melanogaster. Dev. Biol.* **135**, 376-391.

Messing, J. (1983). New M13 vectors for cloning. *Methods Enzymol.* **101**,20-78.

Miller, J., McLachlan, A.D. & Klug, A. (1985). Repetitive zinc-binding domains in the protein transcription factor IIIA from *Xenopus* oocytes. *EMBO J.* **4**, 1609-1614.

Mitrani, E., Ziv, T., Thomsen, G., Shimoni, Y., Melton, D.A. & Bril, A. (1990). Activin can induce the formation of axial structures and is expressed in the hypoblast of the chick. *Cell* **63**, 495-501.

Mlodzik, M., Fjose, A. & Gehring, W.H. (1988). Molecular structure and spatial expression of a homeobox gene from the *labial* region of the Antennapedia complex. *EMBO J.* **7**, 2569-2578.

Morata, G. & Lawrence, P.A. (1975). Control of compartment development by the *engrailed* gene in *Drosophila. Nature* **255**, 614-617.

Morgan, T.H. (1934). *Embryology and genetics*. University Press New York.

Morriss, G. M. (1972). Morphogenesis of the malformations induced in rat embryos by maternal hypervitaminosis A. *J. Anat.* **113**, 241-250.

Morriss, G. M. & Thorogood, P. V. (1978). An approach to cranial neural crest migration and differentiation in mammalian embryos. in *Development in Mammals* vol. **3** (ed. M. H. Johnson), 387-400.

Morriss-Kay, G.M. (1981). Growth and development of pattern in the cranial neural epithelium of rat embryos during neurulation. *J. Embryol. exp. Morph.* **65** (supplement), 225-241.

Mount, S. (1982). A catalogue of splice junction sequences. *Nucleic Acid Res.* **10**, 459-472.

Muller, M.M., Carrasco, A.E. & DeRobertis, E.M. (1984). A homeobox-containing gene expressed during oogenesis in *Xenopus*. *Cell* **39**, 157-162.

Muller-Hill, B., Crapo, L. & Gilbert, W. (1968). Mutants that make more *lac* repressor. *Proc Natl. Acad Sci. U.S.A.* **59**, 1259-1264.

Muller-Immergluck, M., Schaffner, W. & Matthias, P. (1990). Transcription factor Oct-2A contains functionally redundant activating domains and works selectively from a promoter but not from a remote enhancer position in non-lymphoid (HeLa) cells. *EMBO J.* **9**, 1625-1634.

Murphy, P., Davidson, D.R. & Hill, R.E. (1989). Segment specific expression of a homeobox-containing gene in the mouse hindbrain. *Nature* **341**, 156-159.

Murphy, P. & Hill, R. E. (1990). Expression of the mouse *labial*-like homeobox containing genes, *Hox 2.9* and *Hox 1.6*, during segmentation of the hindbrain. *Development* in press.

Nakamura, O. & Matsuzawa, T. (1967). Differentiation capacity of the marginal zone in the morula and blastula of *Triturus pyrrhogaster*. *Embryologia* **9**, 223-237.

Neal, H.V., (1918). Neuromeres and metameres. *J. Morph.* **31**, 293-315.

Nieuwkoop, P.D. (1969). The formation of mesoderm in Urodele amphibians. I. Induction by the endoderm. *Wilhelm Roux' Arch. EntwMech. Org.* **162**, 341-373.

Nieuwkoop, P.D. & others (1952). Activation and organisation of the central nervous system of amphibians. *J. exp. Zool.* **120**, 1-31.

Njolstad, P.R., Molven, A., Hordvik, I., Apold, J. & Fjose, A. (1888). Primary structure, developmentally regulated expression and potential duplication of the zebrafish homeobox gene ZF-21. *Nucl. Acids Res.* **16**, 9097-9111.

Noden, D.M. (1975). An analysis of the migratory behavior of avian cephalic neural crest cells. *Devl Biol.* **42**, 106-130.

Noden, D. M. (1983). The role of the neural crest in patterning of avian cranial skeletal, connective and muscle tissues. *Devl Biol.* **96**, 144-165.

Nornes,H.O., Dressler, G.R., Knapik, E.W., Deutsch, U. & Gruss, P. (1990). Spatially and temporally restricted expression of Pax-2 during murine neurogenesis. *Development* **109**, 797-809.

Norrander, J., Kempe, T. & Messing, J. (1983). Construction of improved M13 vectors using oligodeoxynucleotide-directed mutagenesis. *Gene* **26**, 101-106.

Nusslein-Volhard, C. & Wieschaus, E. (1980). Mutations affecting segment number and polarity in *Drosophila*. *Nature* **287**, 795-801.

Nusslein-Volhard, C., Wieschaus, E. & Kluding, H. (1984). Mutations affecting the pattern of the larval cuticle in *Drosophila melanogaster*. I. Zygotic loci on the second chromosome. *Wilhelm Roux Arch. devl Biol.* **193**, 267-282.

Odenwald, W.F., Taylor, C.F., Palmer-Hill, F.J., Friedrich, V., Tani, M. & Lazzarini, R.A. (1987). Expression of a homeodomain protein in non-contact inhibited cultural cells and post mitotic neurons. *Genes Dev.* **1**, 482-496.

O'Farrell, P. & Scott, M.P. (1986). Spatial programming of gene expression in early *Drosophila* embryogenesis. *Ann. Rev. Cell Biol.* **2**, 49-80.

Oliver, G., Wright, C. V. E. Hardwicke, J. & DeRobertis, E. M. (1988). Differential antero-posterior expression of two proteins encoded by a homeobox gene in *Xenopus* and mouse embryos. *EMBO J.* **7**, 3199-3209.

Orr, H., (1887). Contribution to the embryology of the lizard. *J. Morph.* **1**, 311-373.

Otwinowski, Z., Schevitz, R.W., Zhang, R-G., Lawson, C.L., Joachimiak, A., Marmorstein, R.Q., Luisi, B.F. & Sigler, P.B. (1988). Crystal structure of *trp* repressor/operator complex at atomic resolution. *Nature* **335**, 321-329.

Peifer, M., Karch, F. & Bender, W. (1987). *Genes Dev.* **1**, 891-898.

Perrimon, N., Engstrom, L. & Mahowald, A.P. (1982). The effects of zygotic lethal mutations on female germ line functions in *Drosophila. Devl Biol.* **105**, 404-414.

Petkovich, M., Brand, N. J., Krust, A. & Chambon, P. (1987). A human retinoic acid receptor which belongs to the family of nuclear receptors. *Nature* **330**, 444-450.

Poole, S., Kouvar, L.M., Drees, B. & Kornberg, T. (1985). The engrailed locus of *Drosophila*: structural analysis of an embryonic transcript. *Cell* **40**, 37-43.

Proudfoot, N.J. & Brownlee, G.G. (1976) Non-coding region sequences in eukaryotic messenger RNA. *Nature* **262**, 211-214.

Puschel, A.W., Balling, R. & Gruss, P. (1990). Position-specific activity of the *Hox 1.1* promoter in transgenic mice. *Development* **108**, 435-442.

Qian, Y. Q., Billiter, M., Otting, G., Muller, M., Gehring, W.J. & Wuthrich, K. (1989). The structure of the *Antennapedia* homeodomain by NMR spectroscopy in solution: comparison with prokaryotic repressors. *Cell* **59**, 573-580.

Rebagliati, M.R., Weeks, D.L., Harvey, R.P. & Melton, D.A. (1985). Identification and cloning of localised maternal RNAs from *Xenopus* eggs. *Cell* **42**, 769-777.

Regulski, M., Harding, K., Kostriken, R., Karch, F., Levine, M. & McGinnis, W. (1985). Homeo-box genes of the *Antennapedia* and *bithorax* complexes of *Drosophila. Cell* **43**, 71-80.

Rigby, P.W., Dieckmann, M., Rhodes, M. & Berg, P. (1977). Labelling deoxyribonucleic acid to high specific activity in vitro by nick translation with DNA polymerase I. *J. Mol. Biol.* **113**, 237-251.

Robert, B., Barton, P., Minty, A., Daubas, P., Weydert, A., Bonhomme, F., Catalan, J., Chazottes, D., Guenet, J.-L. & Buckingham, M. (1985). Investigation of genetic linkage between myosin and actin genes using an interspecific mouse backcross. *Nature* **314**, 181-183.

Robert, B., Sassoon, D., Jacq, B., Gehring, W. & Buckingham, M. (1889). Hox 7, a mouse homeobox gene with a novel pattern of expression during embryogenesis. *EMBO J.* **8**, 91-100.

Robertson, M. (1987). Towards a biochemistry of morphogenesis. *Nature* **330**, 420-422.

Rosa, F.M. (1989). *Mix-1*, a homeobox mRNA inducible by mesoderm inducers, is expressed mostly in the presumptive ectodermal cells of *Xenopus* embryos. *Cell* **57**, 965-974.

Rosa, F., Roberts, A.B., Danielpour, D., Dart, L.L., Sporn, M.B. & Dawid, I.B. (1988). Mesoderm induction in amphibians: the role of TGFb2-like factors. *Science* **239**, 783-785.

Rosenberg, U.B., Schroder, C., Preiss, A., Kienlin, A., Cote, S., Riede, I. & Jackle, H. (1986). Structural homology of the product of the *Drosophila Krupple* gene with *Xenopus* transcription factor IIIA. *Nature* **319**, 336-339.

Ruberte, E., Dolle, P., Chambon, P. & Morriss-Kay, G. (1991). Retinoic acid receptors and cellular retinoid binding proteins: II. Their differential pattern of transcription during early morphogenesis in mouse embryos. *Development*, in press.

Ruberte, E., Dolle, P., Krust, A., Zelent, A., Morriss-Kay, G. & Chambon, P. (1990). Specific spatial and temporal distribution of retinoic acid receptor gamma transcripts during mouse embryogenesis. *Development* **108,** 213-222.

Rubin, M.R., King, W., Toth, L.E., Sawzuk, I.S., Levine, M., D'Estachio, P. & Nguyen-Huu, M.C. (1987). Murine *Hox 1.7* homeobox gene: cloning, chromosomal location and expression. *Mol. Cell. Biol.* **7**, 3836-3841.

Rubock, M.J., Larin, Z., Cook, M., Papalopulu, N., Krumlauf, R. & Lehrach, H. (1990). A yeast artificial chromosome containing the mouse homeobox cluster *Hox 2. Proc. Natl. Acad. Sci. USA* **87,** 4751-4755.

Ruddle, F.H., Hart, C.P., Awgulwitsch, A., Fainsod, A., Utset, M., Dalton, D., Kerk, N., Rabin, M., Ferguson-Smith, A., Fienberg, A. & McGinnis, W. (1985). Mammalian homeo-box genes. *CSH Symp. Quant. Biol.* **50**, 277-284.

Rugh, R. (1990). *The Mouse, its reproduction and development*. Oxford University Press.

Ruiz i Altaba, A. & Melton, D.A. (1989a). Bimodal and graded expression of the *Xenopus* homeobox gene *Xhox3* during embryonic development. *Development* **106**, 173-183.

Ruiz i Altaba, A. & Melton, D.A. (1989b). Interractions between peptide growyh factors and homeobox genes in the establishment of antero-posterior polarity of frog embryos. *Nature* **341**, 33-38.

Ruiz i Altaba, A. & Melton, D.A. (1990). Axial patterning and the establishment of polarity in the frog embryo. *TIG* **6,** 57-64.

Saiki, R.K., Gelfard, D.H., Stoffel, S., Scharf, S.J., Higuchi, R., Horn, G.T., Mullis, K.B. & Erlich, H.A. (1988). Primer directed enzymatic amplification of DNA with a thermostable DNA polymerase. *Science* **239**, 487-491.

Sakai, Y. (1987). Neurulation in the mouse I. The ontogenesis of neural segments and the determination of topographical regions in a central nervous system. *The Anatomical Record* **218**, 450-457.

Sambrook, J., Fritsch, E. F. & Maniatis, T. (1989). *Molecular cloning; a laboratory manual.* Second edition. Cold Spring Harbour Laboratory Press.

Sanchez-Herrero, E., Vernos, I., Marco, R. & Morata, G. (1985). Genetic organisation of *Drosophila* bithorax complex. *Nature* **313,** 108-113.

Sanger, F. & Coulson, A.R. (1978). The use of thin acrylamide gels for DNA sequencing. *FEBS Letters* **87**, 107-110.

Sanger, F., Nicklen, S. & Coulson, A.R. (1977). DNA sequencing with chain terminating inhibitors. *Proc. Natl. Acad. Sci. U.S.A.* **74,** 5463-5467.

Satre, M. A. & Kochar, D. M. (1989). Elevations in the endogenous levels of the putative morphogen retinoic acid in embryonic mouse limb buds associated with limb dysmorphogenesis. *Delelopmental Biology* **133,** 529-536.

Sauer, R.T., Yocum, R.R., Dolittle, R.F., Lewis, M. & Pabo, C.O. (1982). Homology among DNA binding proteins suggests use of a conserved supersecondary structure. *Nature* **298,** 447-451.

Saxen, L. (1977). Directive versus permissive induction: a working hypothesis. In *Cell and Tissue Interractions* (ed. J. Lash & M. Burger), p1-9 New York: Raven Press.

Saxen, L. (1987). *Organogenesis of the kidney.* Cambridge Universuty Press.

Schneuwly, S., Kuroiwa, A., Baumgartner, P. & Gehring, W.J. (1986). Structural organisation and sequence of the homeotic gene *Antennapedia* of *Drosophila melanogaster. EMBO J.* **5,** 733-739.

Schughart, K., Utset, M.F., Awgulewitsch, A. & Ruddle, F.H. (1988). Structure and expression of *Hox 2.2,* a murine homeobox-containing gene. *Proc. Natl. Acad. Sci. USA* **85,** 5582-5586.

Schulze, F., Chowdry, K., Zimmer, A., Drescher, U. & Gruss, P. (1987). The murine homeobox gene product, *Hox 1.1* protein, is growth controlled and associated with chromatin. *Differentiation* **36,** 130-137.

Schupbach, T. & Wieschaus, E. (1986). Maternal effect mutations affecting the segmental pattern of *Drosophila. Wilhelm Roux Arch. devl Biol.* **195,** 302-307.

Schwartzberg, P.L., Goff, S.P. & Robertson, E.J. (1989). Germ-line transmission of a *c-abl* mutation produced by targeted gene disruption in ES cells. *Science* **246,** 799-803

Scott, M.P. & Carroll, S.B. (1987). The segmentation and homeotic gene network in early *Drosophila* development. *Cell* **51,** 689-698.

Scott, M.P. & Weiner, A.J. (1984). Structural relationships among genes that control development: sequence homology between *Antennapedia, Ultrabithorax* and *fishi tarazu* loci of *Drosophila. Proc. Natl. Acad. Sci. USA.* **81,** 4115-4119.

Scott, M.P., Weiner, A.J., Hazelrigg, T.I., Polisky, B.A., Pirotta, V., Scalenghe, F. & Kaufmann, T.C. (1983). The molecular organisation of the *Antennapedia* locus of *Drosophila. Cell* **35,** 763-776.

Sharpe, C.R., Fritz, A., DeRobertis, E.M. & Gurdon, J.B. (1987). A homeobox-containing marker of posterior neural differentiation shows the importance of predetermination in neural induction. *Cell* **50,** 749-758.

Sharpe, C.R. & Gurdon, J.B. (1990). The induction of anterior and posterior neural genes in *Xenopus laevis. Development* **109,** 765-774.

Sharpe, P.T., Miller, J.R., Evans, E.P., Burtenshaw, M.D. & Gaunt, S.J. (1988). *Hox 6*: a new homeobox gene locus. *Development* **102,** 397-407.

Shepherd, J.W.C., McGinnis, W., Carrasco, A.E., DeRobertis, E.M. & Gehring, W.J. (1984). Fly and frog homeodomains show homologies with yeast amting regulatory proteins. *Nature* **310,** 70-71.

Shuttleworth, J. & Colman, A. (1988). Antisense oligonucleotide-directed cleavage of mRNA in *Xenopus* oocytes and eggs. *EMBO J.* **7,** 427-434.

Simeone, A., Acampora, D., Arcioni, L., Andrews, P. W., Boncinelli, E. & Mavilio, F. (1990). Sequential activation of *HOX 2* homeobox genes by retinoic acid in human embryonal carcinoma cells. *Nature* **346,** 763-766.

Sive, H. L., Draper, B. W., Harland, R. M. & Weintraub, H. (1990). Identification of a retinoic acid-sensitive period during primary axis formation in *Xenopus laevis. Genes & Development* **4,** 932-942.

Slack, J.M.W. (1983). *From egg to embryo.* Cambridge University Press.

Slack, J.M.W. (1987). Morphogenetic gradients- past and present. *TIBS* **12,** 200-204.

Slack, J.M.W., Darlington, B.G., Heath, J.K. & Godsave, S.F. (1987). Mesoderm induction in early *Xenopus* embryos by heparin binding growth factors. *Nature* 326, 197-200.

Slack, J.M.W. & Forman, D. (1980). An interraction between dorsal and ventral regions of the marginal zone in early amphibian embryos. *J. Embryol. exp. Morph.* **56,** 283-299.

Slack, J.M.W. & Isaacs, H. (1989). Presence of basic fibroblast growth factor in the early *Xenopus* embryo. *Development* **105,** 147-153.

Slack, J.M.W., Isaacs, H.V. & Darlington, B.G. (1988). Inductive effects of fibroblast growth factor and lithium ion on *Xenopus* blastula ectoderm. *Development* 103, 581-590.

Smith, J.C. (1987). A mesoderm-inducing factor is produced by a *Xenopus* cell line. *Development* **99,** 3-14.

Smith, J.C. (1989). Mesoderm induction and mesoderm inducing factors in early amphibian development. *Development* **105,** 665-677.

Smith, J.C., Price, B.M.J., Van Nimmen, K. & Huylebroeck, D. (1990). Identification of a potent *Xenopus* mesoderm-inducing factor as a homologue of activin A. *Nature* **345,** 729-731.

Smith, J.C. & Slack, J.M.W. (1983). Dorsalisation and neural induction: properties of the organiser in *Xenopus laevis. J. Embryol. exp. Morph.* **78,** 299-317.

Smith, J.C., Yaqoob, M. & Symes, K. (1988). Purification, partial characterisation and biological effects of XTC mesoderm-inducing factor. *Development* **103,** 591-600.

Snow, M.L. (1981). Autonomous development of parts isolated from primitive-streak-stage mouse embryos. Is development clonal? *J. Embryol. exp. Morph.* **65,** 269-287.

Sokol, S., Wong, G. & Melton, D. (1990). A mouse macrophage factor induces head structures and organises a body axis in *Xenopus. Science* **249,** 561-564.

Southern, E.M. (1975). Detection of specific sequences among DNA fragments seperated by gel electrophoresis. *J. Mol. Biol.* **98,** 503-517.

Sternberg, P.W. & Horvitz, H.R. (1986). Pattern formation during vulval development in *C. elegans. Cell* **44,** 761-772.

Steward, R., Zusman, S.B., Huang, L.H. & Scheldl, P. (1988). The *dorsal* protein is distributed in a gradient in early *Drosophila* embryos. *Cell* **55,** 487-495.

Streeter, G.L. (1908). The nuclei of origin of the cranial nerves in the 10 MM. human embryo. *Anat. Rec.* **2,** 111-115.

Strickland, S. & Mahdavi, V. (1978). The induction of differentiation in teratocarcinoma stem cells by retinoic acid. *Cell* **15,** 393-403.

Strickland, S., Smith, K.K., & Marotti, K.R. (1980). Hormonal induction of differentiation in teratocarcinoma stem cells: generation of parietal endoderm by retinoic acid and dibutyryl cAMP. *Cell* **21,** 347-355.

Struhl, G. (1981). A gene product required for the correct initiation of segment determination in *Drosophila. Nature* **293,** 36-41.

Struhl, G. (1983). Role of the *esc* gene product in ensuring the selective expression of segment specific homeotic genes in *Drosophila. J. Embryol. exp. Morph.* **76,** 297-331.

Struhl, G., Struhl, K. & Macdonald, P.M. (1989). The gradient morphogen *Bicoid* is a concentration dependent transcriptional activator. *Cell* **57,** 1259-1273.

Struhl, G. & White, R.A.H. (1985). Regulation of the *Ultrabithorax* gene of *Drosophila* by other bithorax complex genes. *Cell* **43,** 507-519.

Sundarwati, S. & Nieuwkoop, P.D. (1971). Mesoderm formation in the Anuran *Xenopus laevis* (Daudin). *Wilhelm Roux' Arch. EntwMech. Org.* **166,** 189-204.

Sundin, O.H., Busse, H.G., Rogers, M.A., Gudas, L.J. & Eichele, G. (1990). 'Region-specific expression in early chick and mouse embryos of *Ghox-lab* and *Hox 1.6,* vertebrate homeobox-containing genes relateed to *Drosophila labial.' Development* **108,** 47-58.

Sundin, O. H. & Eichele, G. (1990). A homeo domain protein reveals the metameric nature of the developing chick hindbrain. *Genes & Development* **4,** 1267-1276.

Tam, P.P.L. & Snow, M.H.L. (1980). The in vitro culture of primitive streak stage mouse embryos. *J. Embryol. Exp. Morphol.* **59,** 131-143.

Tam, P.P.L., Meier, S. & Jacobson, A.G. (1982). Differentiation of the metameric pattern in the embryonic axis of the mouse. II. Somitomeric organisation of the presomitic mesoderm. *Differentiation* **21,** 109-122.

Tan, S.S. & Morriss-Kay, G.M. (1986). Analysis of cranial neural crest cell migration and early fates in postimplantation rat chimaeras. *JEEM* **98,** 21-58.

Tannahill, D. & Melton, D.A. (1989). Localised synthesis of the *Vg1* protein during early *Xenopus* development. *Development* **106,** 775-785.

Tartof, K.D. & Hobbs, C.A. (1987). Improved media for growing plasmid and cosmid clones. *Focus (Bethesda Research Laboratories Inc.)* **9:2,** 12.

Tautz, D. (1988). Regulation of the *Drosophila* segmentation gene *hunchback* by two maternal morphogenetic centres. *Nature* **332,** 281-284.

Tautz, D., Lehmann, R., Schnurch, H., Schuh, R., Seifert, E., Kienlin, A., Jones, K. & Jackle, H. (1987). Finger protein of novel structure encoded by *hunchback,* a second member of the gap class of *Drosophila* segmentation genes. *Nature* **327,** 383-389.

Taylor, B.A. (1982). Recombinant inbred strains, p. 423-438. In H.C. Morse III (ed), *Origins of inbred mice.* Academic Press, Inc., New York.

Tessier-Lavigne, M., Placzek, M., Lumsden, A.G.S. Dodd, J. & Jessell, T.M. (1988). Chemotrophic guidance of developing axons in the mammalian central nervous system. *Nature* **336**, 775-778.

Tickle, C. A., Summerbell, D. & Wolpert, L. (1975). Positional signalling and specification of digits in chick limb morphogenesis. *Nature* **254**, 199-202.

Tickle, C., Alberts, B., Wolpert, L. & Lee, J. (1982). Local application of retinoic acid to the limb bud mimics the action of the polarising region. *Nature* **296**, 564-566.

Thaller, C. & Eichele, G. (1987). Identification and spatial distribution of retinoids in the developing chick limb bud. *Nature* **327**, 625-628.

Thaller, C. & Eichele, G. (1990). Isolation of 3,4-didehydroretinoic acid, a novel morphogenetic signal in the chick wing bud. *Nature* **345**, 815-819.

Theiler, K. (1972). *The House Mouse*. Springer-Verlag.

Thomas, K.R. & Capecchi, M.R. (1990). Targeted disruption of the murine int-1 proto-oncogene resulting in severe abnormalities in midbrain and cerebellar development. *Nature* **346**, 847-850.

Thompson, S., Clarke, A.R., Pow, A.M., Hooper, M.L. & Melton, D.W. (1989). Germ-line transmission and expression of a corrected HPRT gene produced by gene targeting in embryonic stem cells. *Cell* **56**, 313-321.

Thomsen, G., Woolf, T., Whitman, M., Sokol, S., Vaughan, J., Vale, W. & Melton, D.A. (1990). Activins are expressed early in *Xenopus* embryogenesis and can induce axial mesoderm and anterior structures. *Cell* **63**, 485-493.

Toth, L.E., Slawin, K.L., Pintar, J.E. & Nguyen-Huu, M. (1987). Region specific expression of mouse homeobox genes in the embryonic mesoderm and central nervous system. *Proc. Natl. Acad. Sci. USA* **84**, 6790-6794.

Treisman, J., Gonczy, P., Vashishtha, M., Harris, E. & Desplan, C. (1989). A single amino acid can determine the DNA binding specificity of homedomain proteins. *Cell* **59**, 553-562.

Tuckett, F., Lim, L. & Morriss-Kay, G.M. (1985). The ontogenesis of cranial neuromeres in the rat embryo I. A scanning electron microscope and kinetic study. *J. Embryol. exp. Morph.* **87**, 215-228.

Tuckett, F. & Morriss-Kay, G.M. (1985). The ontogenesis of cranial neuromeres in the rat embryo II. A transmission electron microscope study. *J. Embryol. exp. Morph.* **88**, 231-247.

Utset, M.F., Awgulewitsch, A., Ruddle, F.H. & McGinnis, W. (1987). Region specific expression of two mouse homeobox genes. *Science* **235**, 1379-1382.

Vaage, S. (1969). The segmentation of the primitive neural tube in chick embryos. *Adv. Anat. Embryol. Cell Biol.* **41**, 1-88.

Vale, W., Rivier, J., Vaughan, J., McClintock, R., Corrigan, A., Wilson, W., Karr, D. & Spiess, J. (1986). Purification and characterisation of an FSH releasing protein from porcine ovarian follicular fluid. *Nature* **321**, 776-779.

Vieira, J. & Messing, J. (1982). The pUC plasmids an M13mp7-derived system for insertion mutagenesis and sequencing with synthetic universal primers. *Gene* **19**, 259-269.

Wagner, M., Thaller, C., Jessell, T. & Eichele, G. (1990). Polarising activity and retinoid synthesis in the floor plate of the neural tube. *Nature* **345**, 819-822.

Wang, S.Y. & Gudas, L.J. (1983). Isolation of cDNA clones specific for collagen IV and laminin from mouse teratocarcinoma cells. *Proc. Natl. Acad. Sci. USA* **80**, 5880-5884.

Wang, S.Y., LaRosa, G.J. & Gudas, L.J. (1985). Molecular cloning of gene sequences transcriptionally regulated by retinoic acid dibutyryl cyclic AMP in cultured mouse teratocarcinoma cells. *Devl Biol.* **107**, 75-86.

Warner, A. & Gurdon. J.B. (1987). Functional gap junctions are not required for muscle gene activation by induction in *Xenopus* embryos. *J. Cell Biol.* **104**, 557-564.

Weber, I.T., McKay, D.B. & Steutz, T.A. (1982). Two helix DNA binding motif of CAP found in lac repressor and gal repressor. *Nucleic Acids Res.* **10**, 5085-5102.

Webster, W. S., Johnston, M.C., Lammer, E. J. & Sulik, K. K. (1986). Isotretinoin embryopathy and the cranial neural crest: an in vivo and in vitro study. *J. of Craniofacial Genetics and Developmental Biology* **6**, 211-222.

Wedden, S.E., Pang, K. & Eichele, G. (1989). Expression pattern of homeobox containing genes during chick embryogenesis. *Development* **105**, 639-650.

Wedden, S. E., Ralphs, J. R. & Tickle, C. (1988). Pattern formation in the facial primordia. *Development* **103** *Supplement,* 31-40.

Weeks, D.L. & Melton, D.A. (1987). A maternal mRNA localised to the vegetal hemisphere in *Xenopus* eggs codes for a growth factor related to TGFb. *Cell* **51**,861-867.

Weir, M.P. & Kornberg, T. (1985). Patterns of *engrailed* and *fushi tarazu* transcripts reveal novel intermediate stages in *Drosophila* segmentation. *Nature* **318**, 433-439.

Wessells, N.K. (1977). *Tissue interractions and development.* Calif. usa: W.A. Benjamin.

Wharton, R.P. & Ptashne, M. (1987). A new specificity mutant of 434 repressor that defines an amino acid base-pair contact. *Nature* **326**, 888-891.

White, R. & Wilcox, M. (1985). The distribution of *Ultrabithorax* proteins in *Drosophila. EMBO J.* **4**, 2035-2043.

Wieschaus, E., Nusslein-Volhard, C. & Jurgens, G. (1984). Mutations affecting the pattern of the larval cuticle in *Drosophila melanogaster.* I. Zygotic loci on the X chromosome and the fourth chromosome. *Wilhelm Roux Arch. devl Biol.* **193**, 267-282.

Wilkinson, D.G., Bailes, J.A. & McMahon, A.P. (1987). Expression of the proto-oncogene *int*-1 is restricted to specific neural cells in the developing mouse embryo. *Cell* **50**, 79-88.

Wilkinson, D.G., Bhatt, S., Chavrier, P., Bravo, R., Charnay, P. (1989a). Segment specific expression of a zinc finger gene in the developing nervous system of the mouse. *Nature* **337**, 461-464.

Wilkinson, D.G., Bhatt, S., Cook, M., Boncinelli, E. & Krumlauf, R. (1989b). Segmental expression of *Hox 2* homeobox genes in the developing mouse hindbrain. *Nature* **341**, 405-409.

Wilkinson, D.G., Peters, G., Dickson, C. & McMahon, A.P. (1988). Expression of the FGF-related proto-oncogene *int-2* during gastrulation and neurulation in the miuse. *EMBO J* **7**, 691-695.

Winslow, G.M., Hayashi, S., Krasnow, M., Hogness, D.S. & Scott, M.P. (1989). Transcriptional activation by *Antennapedia* and *fushi tarazu* proteins in cultured *Drosophila* cells. *Cell* **57**, 1017-1030.

Wolgemuth, D.J., Behringer, R.B., Mostoller, M.P., Brinster, R.L. & Palmiter, R.D. (1989. Transgenic mice overexpressing the mouse homeobox-containing gene Hox 1.4 exhibit abnormal gut development. Nature 337, 464-467.

Wolgemuth, D.J., Viviano, C.M., Gizang-Ginsberg, E., Frohman, M.A., Joyner, A.L. & Martin, G.R. (1987). Differential expression of the mouse homeobox containing gene *Hox 1.4* during male germ cell differentiation and embryonic development. *Proc. Natl. Acad. Sci. USA* **84**, 5813-5817.

Wolpert, L. (1969). Positional information and the spatial pattern of cellular differentiation. *J. Theoret. Biol.* **25**, 1-47.

Woodland, H.R. (1989). Mesoderm formation in *Xenopus. Cell* **59,** 767-770.

Woychik, R.P., Maas, R.L., Zeller, R., Vogt, T.F. & Leder, P. (1990). 'Formins': Proteins deduced from the alternative transcripts of the *limb deformity* gene. *Nature* **346**, 850-853.

Wright, C.V.E., Cho, K.W.Y., Hardwicke, J., Collins, R.H. & DeRobertis, E. (1989). Interference with function of a homeobox gene in *Xenopus* embryos produces malformations of the anterior spinal cord. *Cell* **59**, 81-93.

Yanisch-Peron, C., Vieria, J. & Messing, J. (1985). Improved M13 phage cloning vectors and host strains: nucleotide sequences of the M13mp18 and pUC19 vectors. *Gene* **33**, 103-119.

Zakany, J., Tuggle, C.K., Patel, M.D. & Nguyen-Huu, M.C. (1988). Spatial regulation of homeobox gene fusions in the embrtonic central nervous system of transgenic mice. *Neuron* **1**, 679-691.

Zelent, A., Krust, A., Petkovich, M., Kastner, P. & Chambon, P. (1989). Cloning of murine a and b retinoic acid receptors and a novel receptor g predominantly expressed in skin. *Nature* **339,** 714-717.

Zijlstra, M., Li, E., Sajjadi, S., Subramani, S. & Jaenisch, R. (1989). Germline transmission of a disrupted ß2 microglobulin gene produced by homologous recombination in embryonic stem cells. *Nature* **342**, 435-438.

Zimmer, A. & Gruss, P. (1989). Production of chimaeric mice containing embryonic stem (ES) cells carrying a homeobox *Hox 1.1* allele mutated by homologous recombination. *Nature* **338**, 150-154.

Appendix 1

Abbreviations

AP	anterioposterior
bp	base pair
cDNA	complimentary DNA
cm	centimetre
CNS	central nervous system
CRABP	cellular retinoic acid binding protein
CRBP	cellular retinol binding protein
C-terminal	carboxy-terminal
dATP	deoxy adenine triphosphate
dCTP	deoxy cytosine triphosphate
dGTP	deoxy guanine triphosphate
dTTP	deoxy thymine triphosphate
dNTP	deoxy nucleoside triphosphate
DNA	deoxynucleicacid
DV	dorsoventral
FGF	fibroblast growth factor
hr	hour
HSV-1-TK	herpes simplex virus 1 thymidine kinase
ICM	inner cell mass
kb	kilobase
kD	kiloDalton
kg	kilogram
krpm	kilo-revolutions per minute
l	litre
M	molar
MIF	mesoderm inducing factor
min	minute
ml	mililitre
mM	milimolar
ng	nanogram
N-terminal	amino-terminal
^{32}P	phosphate isotope, molecular weight = 32
PCR	polymerase chain reaction

RA	retinoic acid
RAR	retinoic acid receptor
RNA	ribonucleic acid
^{35}S	sulphur isotope, molecular weight = 35
TGFß	transforming growth factor type ß
UV	ultraviolet
+ve	positive
-ve	negative
ZPA	zone of polarising activity
^{o}C	degrees celcius
µl	microlitre
µm	micrometre
µM	micromolar
w/v	weight per volume

cence image of decorated actin filaments before ATP addition, and electron micrographs of the corresponding field. Filaments were well preserved and remained at the same locations despite the incubation with antibodies and the negative staining. More than 99% of the filaments were identified in both images (left). At a higher magnification (right), gold patches were seen to align along the actin filaments. The density of gold particles in the background reflects a random distribution of HMM which was not of sufficient density to support actin filament movement. Figure 2*b* shows a fluorescence image before ATP addition, and electron micrographs of the corresponding area after actin filaments had moved off the tracks in the presence of ATP. Gold particles still remained along the initial tracks of decorated filaments (right). The width of the track or the density of gold particles in the background did not increase appreciably. Furthermore, the density of the gold particles along the tracks remained almost the same (31.1 ± 2.8 particles per μm, $n = 20$ filaments, before ATP addition; 30.6 ± 3.4 particles per μm, $n = 23$, after ATP addition). These observations indicate that the HMM molecules associated with the nitrocellulose surface while maintaining unipolar binding to actin filaments, and stayed attached to the surface, probably in their original locations, even after actin filaments moved away.

To examine the flexibility of the nitrocellulose-bound HMM in binding to actin filaments, tracks of HMM were formed using long actin filaments, which were subsequently removed by ATP addition, and then very short actin filaments prepared with severin[7] were introduced in the absence of ATP. If HMM could bind to actin filaments lying in either direction along the track, then it was anticipated that we should be able to see arrowhead patterns pointing in opposite directions on a single track of HMM. Figure 3*b* shows several short repeats of arrowheads formed on a long track; they indeed show opposite polarities (see long arrows). Immunogold labelling of severin (Fig. 3, short arrows), which binds to the barbed end of actin filaments[8], corroborates the polarity.

These experiments show that (1) actin filaments can follow the path pre-determined by HMM locations irrespective of the filament length, (2) actin filaments can move along the same HMM track in opposite directions, and (3) HMM bound to a nitrocellulose surface is flexible and can form a rigor complex even with an actin filament having opposite polarity to that of the initial filament. Two earlier observations showed that there is considerable flexibility in the HMM portion of the myosin molecule. First, two heads of HMM bind to actin molecules on the same strand in the absence of ATP[9], showing that the two heads are translationally related in this bound configuration. Second, a single IgG molecule sometimes cross-links two heads of myosin[10,11], showing that the two heads can be arranged so that the same epitopes on the two heads face each other. Thus, there must be a considerable rotational freedom in the head-rod junction or head itself. The head–rod junction is known to have a large degree of freedom (for example, ref. 12); the myosin head may be rotating rapidly around its long axis looking for the correct binding site on the actin. One possible interpretation for the bidirectional movement described here is, therefore, that 180° rotational freedom exists in HMM. This interpretation is strongly supported by the complementary experiments reported by Reedy *et al.*[13]. They showed that myosin heads tethered in a single thick-filament of a mutated *Drosophila* muscle sarcomere can bind with opposite rigor-cross-bridge angles to flanking thin filaments, which are apparently of opposite polarities. □

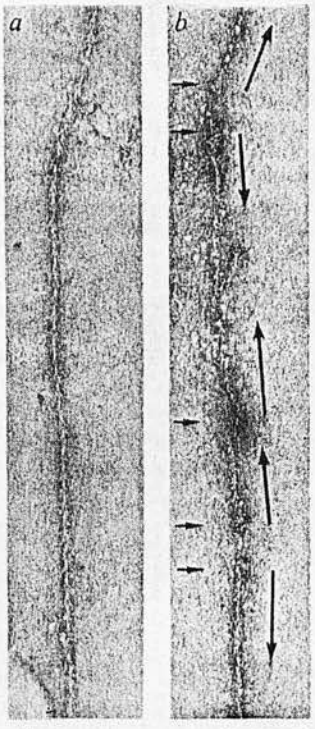


FIG. 3 *a*, Negatively stained actin filament decorated with HMM on a nitrocellulose film of a flow cell; well-ordered arrowhead patterns show unipolar binding of HMM along the actin filament. *b*, Short decorated actin filaments pointing in opposite directions (long arrows) on a single long track of HMM, showing that the HMM bound to the nitrocellulose film could bind to actin filaments with either polarity. HMM tracks were first formed by placing onto the nitrocellulose long decorated actin filaments, which were removed later by addition of ATP; short actin filaments prepared with severin were then introduced in the absence of ATP to form rigor complexes with the HMM on the track. The gold particles (short arrows) show the presence of severin at the barbed ends of the short actin filaments, corroborating the polarity indicated by the arrowhead pattern. Scale bar, 0.2 μm.
METHODS. Short actin filaments were prepared with severin, an actin-filament severing protein[7]. Severin was first mixed with actin labelled with rhodamine-phalloidin at a 1:100 molar ratio in the absence of Ca^{2+}; 0.1 mM Ca^{2+} was subsequently added to the mixture to activate the severin[7]. Antibody against severin and gold-conjugated secondary antibody were applied after introduction of the actin filaments into a flow cell, and the specimens were negatively stained with 1% uranyl acetate.

Received 10 July; accepted 9 August 1989.

1. Yano, M., Yamamoto, Y. & Shimizu, H. *Nature* **299**, 557-559 (1982).
2. Hynes, T. R., Block, S. M., White, B. T. & Spudich, J. A. *Cell* **48**, 953-963 (1987).
3. Toyoshima, Y. Y. *et al. Nature* **328**, 536-539 (1987).
4. Kishino, A. & Yanagida, T. *Nature* **334**, 74-76 (1988).
5. Huxley, H. E. *J. molec. Biol.* **7**, 281-308 (1963).
6. Moore, P. B., Huxley, H. E. & DeRosier, D. J. *J. molec. Biol.* **50**, 279-295 (1970).
7. Yamamoto, K., Pardee, J. D., Reidler, J., Stryer, L. & Spudich, J. A. *J. Cell Biol.* **95**, 711-719 (1982).
8. Giffard, R. G., Weeds, A. G. & Spudich, J. A. *J. Cell Biol.* **98**, 1796-1803 (1984).
9. Craig, R. *et al. J. molec. Biol.* **140**, 35-55 (1980).
10. Winkelmann, D. A. & Lowey, S. *J. molec. Biol.* **188**, 595-612 (1986).
11. Miyanishi, T., Toyoshima, C., Wakabayashi, T. & Matsuda, G. *et al. J. Biochem.*, Tokyo **103**, 458-462 (1988).
12. Kinosita, K., Ishiwata, S., Yoshimura, H., Asai, H. & Ikegami, A. *Biochemistry* **23**, 5963-5975 (1984).
13. Reedy, M. C., Beall, C. & Fyrberg, E. *Nature* **339**, 481-483 (1989).
14. Kron, S. J. & Spudich, J. A. *Proc. natn. Acad. Sci. U.S.A.* **83**, 6272-6276 (1986).

ACKNOWLEDGEMENTS. We thank Dr K. Sutoh, University of Tokyo, for the gift of SI antibodies. This work was supported by a grant from NIH to J.A.S.; Y.Y.T. and C.T. were Postdoctoral Fellows of the Muscular Dystrophy Association.

Segment-specific expression of a homoeobox-containing gene in the mouse hindbrain

Paula Murphy, Duncan R. Davidson & Robert E. Hill

MRC Human Genetics Unit, Western General Hospital, Crewe Road, Edinburgh EH4 2XU, UK

THE process of segmentation, in which the developing embryo is divided into repetitive structures along its antero–posterior (A–P) axis, as a means of organizing and coordinating the body plan is found in a wide range of organisms. In *Drosophila*, homoeotic genes are involved in all levels of segmental organization and in determining segment identity[1]. The roles of these genes in segmentation have been found mainly by mutational studies[2,3], but also by *in situ* hybridization, which has shown their domains of expression. In contrast to *Drosophila*, however, embryonic

expression of homoeobox-containing genes in vertebrate organisms has not been found to follow a segmental pattern[4]. Vertebrate segmentation can be clearly seen in the mesodermal somites, but repetitive morphological structures in the central nervous system (neuromeres)[5–13] have only recently been shown to have developmental significance. Neuromeres in the hindbrain (rhombomeres) have been defined as segmental units by their pattern of nerve formation in the developing chick[14] and by the alternating expression of *Krox-20*, a gene encoding a zinc-finger DNA-binding protein, in the 9.5-day-old mouse[15]. Here we report that a mouse homoeobox-containing gene, *Hox-2.9*, is expressed in a segment-specific manner in the developing mouse hindbrain. This expression is in a region which is flanked by the regions of expression of *Krox-20*, and is precisely contained within a single neuromere, rhombomere 4.

A homoeobox-containing gene mapping to the *Hox-2* gene cluster on chromosome 11 (refs 16 and 17), which we have characterized as *Hox-2.9* (P.M. *et al.*, manuscript in preparation; R. Krumlauf, personal communication), was isolated from an 8.5-day-old mouse embryonic complementary DNA library. We determined the pattern of *Hox-2.9* expression during development by *in situ* hybridization using a probe specific for *Hox-2.9*, 3′ to the homoeobox, and analysing serial sections from 6.5-day-, 8.5-day-, 9.5-day-, 10-day- and 13.5-day-old embryos, using at least four embryos from each stage.

Hox-2.9 transcripts were not detected in the 6.5-day-old embryo, when gastrulation begins. At 8.5 days of development the embryo is at the neural fold stage. At this stage, expression of *Hox-2.9* is detected in both the neural ectoderm and the underlying mesoderm (Fig. 1*a*, *b*). The domain of expression is extensive, with an anterior limit in the developing hindbrain posterior to the preotic sulcus (Fig. 1*c*, *d*). The low level of expression and the absence of morphological markers in the hindbrain, however, made it difficult to locate the anterior limit of expression more precisely. Mesodermal expression was restricted to lateral and presomitic mesoderm; no label was detectable in the somites.

By 9.5 days the expression pattern had changed dramatically, as had the complexity of the embryo. At this stage the segmental units of the hindbrain, the rhombomeres, could easily be seen as a series of neuroepithelial swellings in frontal and parasagittal sections (Figs 2 and 3*d*). Individual rhombomeres could be

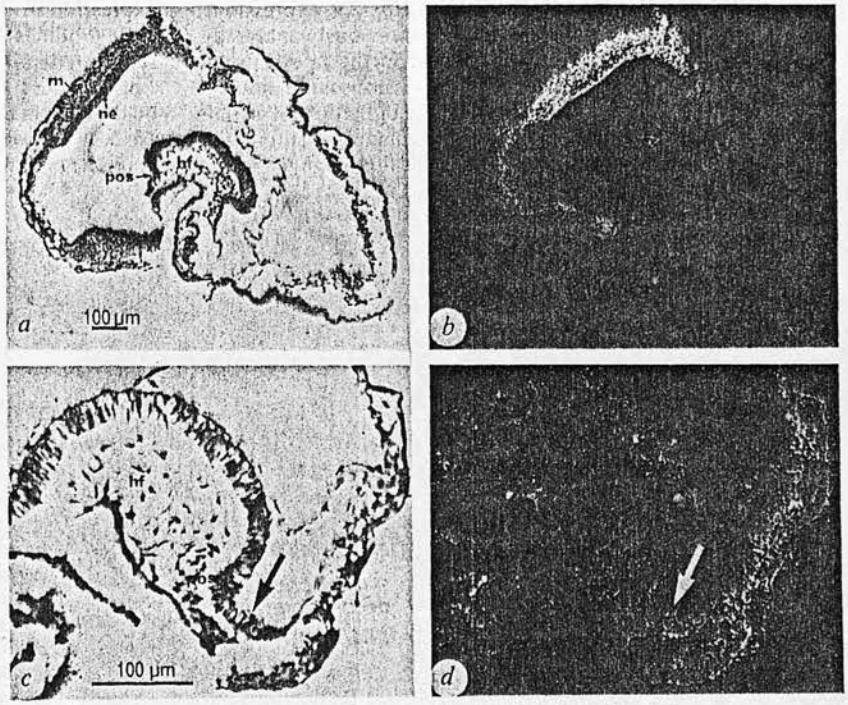


FIG. 1 *In situ* hybridization of a ^{35}S-labelled antisense probe for *Hox-2.9* to 8.5-day-old mouse-embryo sections. *a*, *b*, Sagittal section showing expression in the mesoderm (m) and neural ectoderm (ne). *c*, *d*, Sagittal section through the headfold (hf) region showing the anterior limit of expression in the neuroectoderm and lateral mesoderm (arrow). *a* and *c*, Brightfield photographs; *b* and *d* are the corresponding darkfield photographs. s, Somite; pos, preotic sulcus.
METHODS. The *In situ* hybridization procedure used was as previously described[31], with ~4 × 10^5 d.p.m. of ^{35}S-labelled UTP incorporated into RNA applied per section. Sense-strand control probe was included routinely and showed no hybridization signal above background (data not shown).

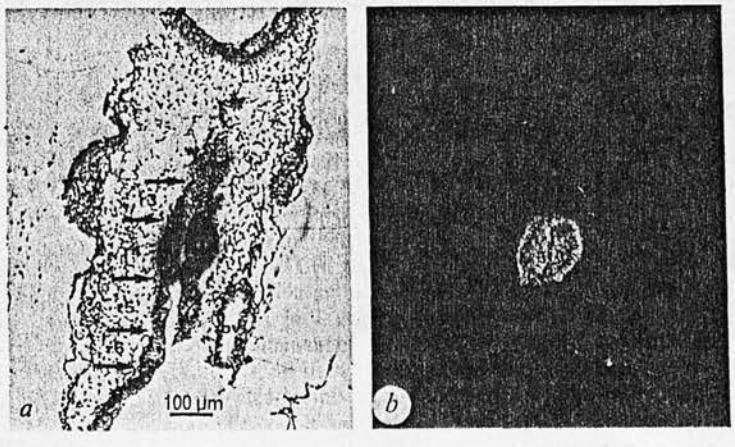


FIG. 2 Frontal section through the hindbrain of a 9.5-day-old embryo showing localized expression of *Hox-2.9* within rhombomere 4 (r4). *a*, Brightfield photograph. *b*, Corresponding darkfield photograph. ov, Otic vesicle; arrows indicate neuromere boundaries; rhombomeres are numbered r3, r4, r5, r6.

identified visually by their position relative to the otic vesicle and numbered as shown in Fig. 4. In the hindbrain, *Hox-2.9* was expressed at this stage in a single segment, rhombomere 4. It was difficult, however, to relate rhombomere 4 to the hindbrain expression at day 8.5, when rhombomeres are not morphologically detectable. Expression was detected throughout rhombomere 4, with the exception of the thin roof plate. The boundaries of expression corresponded precisely with the morphological boundaries of rhombomere 4 (Figs 2 and 3*d*), and Fig. 3*c* shows the precision of this boundary at the cellular level. Rhombomere 4 seems to be the principal site of expression of *Hox-2.9*, because the very localized signal seen at this site was consistently much higher than any other signal detected.

Outside the hindbrain the anterior neural tube was unlabelled at 9.5 days, but expression was detected in the neural tube posterior to the forelimb buds (Figs 3*a*, *b*). Expression was still detected in the presomitic mesoderm but not in the somites (Figs 3*e*, *f*). Gut-associated mesenchyme and the epithelium of the upper gut posterior to the third branchial arch were also labelled (Fig. 3*a*, *b*), the gut-associated mesenchyme having been derived from lateral-plate mesoderm, which showed expression at day 8.5. A low level of expression was detected in the nephrogenic duct (Fig. 3*a*, *b*). Analysis of sections through 10-day-old embryos showed no further change in *Hox-2.9* expression. Unlike most homoeobox-containing genes previously analysed, the expression of *Hox-2.9* did not persist in the late embryo stage: at 13.5 days of development, expression could not be detected in any part of the embryo by examination of serial sections in sagittal, transverse or frontal planes. Neuromeric boundaries could also no longer be seen at this stage.

The expression of *Hox-2.9* in rhombomere 4 indicates that this gene plays a part in the formation or differentiation of this segment. Neuromeres have segmental identities, as seen by the exit points of cranial ganglia from specific neuromeres[14] with the facial and acoustic ganglia extending from rhombomere 4. Independent evidence for the genetic control of development within this neuromere comes from the phentotype of the mouse developmental mutant *kreisler*, which displays faulty segmentation of the hindbrain and specific degeneration of the cells of rhombomere 4 by day 9.5 (ref. 18). The mutation maps to mouse chromosome 2 (ref. 19) and is therefore not allelic with *Hox-2.9*. It is possible that *Hox-2.9* and *kreisler* are both genetic regulators of segmentation in the hindbrain. We can now investigate the relationship between these two genes.

Krox-20 is expressed in rhombomeres 3 and 5 (ref. 15), which flank the region of expression of *Hox-2.9* at day 9.5. At day 8.5, before the neuromeres become visible, *Krox-20* expression is first detected in the region of two neuroepithelial invaginations, proneuromeres-A and -B, indicating that segments start to become established at this stage[15]. *Krox-20* is therefore a likely candidate for involvement in the process of segmentation because its expression precedes the appearance of morphological segments. The spatial and temporal correlation between the expression of *Hox-2.9* and *Krox-20* raises the possibility that their products interact, directly or indirectly, in the process of segmentation in this region of the central nervous system (CNS). Both of these genes are thought to encode regulatory proteins on the basis of the DNA-binding motifs that the proteins possess: *Hox-2.9* encodes a protein that has a helix-turn-helix motif[20] and *Krox-20* encodes a protein with a zinc-finger DNA-binding

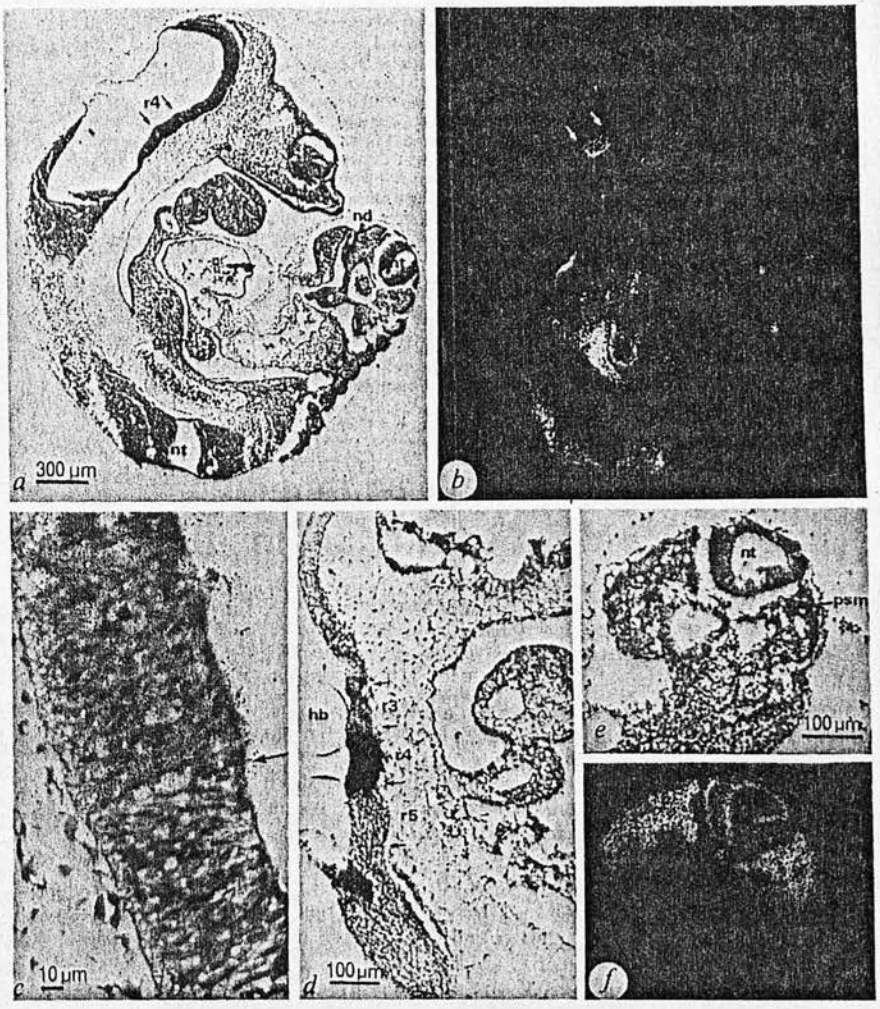


FIG. 3 Expression of *Hox-2.9* in 9.5–10-day-old mouse embryos. *a*, *b*, Sagittal section of a 10-day-old embryo showing expression in the hindbrain (r4, rhombomere 4), the gut-associated mesenchyme (gam), the nephrogenic duct (nd), and the posterior neural tube (nt). *c*, Boundary between r3 and r4 (marked with arrow) in a 10-day sagittal section showing expression of *Hox-2.9* within r4. The short autoradiographic exposure time (two weeks) allowed the labelling of individual cells to be observed and emphasized the sharpness of this expression boundary. *d*, Sagittal 9.5-day embryo section probed with the *Hox-2.9*-specific probe and exposed for a long period (5 weeks) shows the morphology of the hindbrain and the expression of *Hox-2.9* in r4. hb, Hindbrain. *e*, *f*, Transverse section showing expression of *Hox-2.9* in presomitic mesoderm (psm) and the neural tube (nt) in posterior regions of the 9.5-day-old embryo. *a* and *e*, Bright field photographs. *b*, *f*, Corresponding darkfield photographs.

motif[21,22]. There is evidence of interaction between proteins with these regulatory domains within the *Drosophila* segmentation system[23].

An evolutionary relationship has been suggested between genes of the *Hox-2* and *Hox-1* clusters in the mouse and the homoeotic genes of the Antennapedia complex in *Drosophila*[16,17]. The sequence of the *Hox-2.9* homoeobox (data not shown) shows a high degree of similarity to *Hox-1.6*, which is most similar to the *Drosophila* gene *labial*[24]. It is of interest that both *Hox 2.9* and *Hox 1.6* have the same temporal pattern of expression[25] (P.M. *et al.*, manuscript in preparation). *Labial* has a homoeotic role in determining segmental identity and its pattern of expression marks a single ancestral segment in the developing *Drosophila* head[26,27]. Most evidence suggests that segmentation arose independently in invertebrates and vertebrates[28]; the expression of *labial* and *Hox-2.9* in single segments may reflect parallels in the mechanisms that determine the identity of anterior segments in invertebrates and vertebrates.

The results presented here raise the possibility that other homoeobox-containing genes are involved in specifying neural segments. Numerous expression studies on mouse homoeobox-containing genes show that their transcripts are most abundant within the developing CNS, where they occupy large overlapping domains[4]. Several of these transcripts have anterior boundaries to their expression in the hindbrain[16], and our work indicates that studies should now be undertaken to identify these boundaries in sections where neuromeres are well defined. The distinct anterior limit of *Hox-1.5* expression lies at a neuromere boundary[29]. We infer, from its position rostral to the otic vesicle[30], that this boundary lies between rhombomeres 4 and 5, which coincides with the posterior limit of *Hox-2.9* expression.

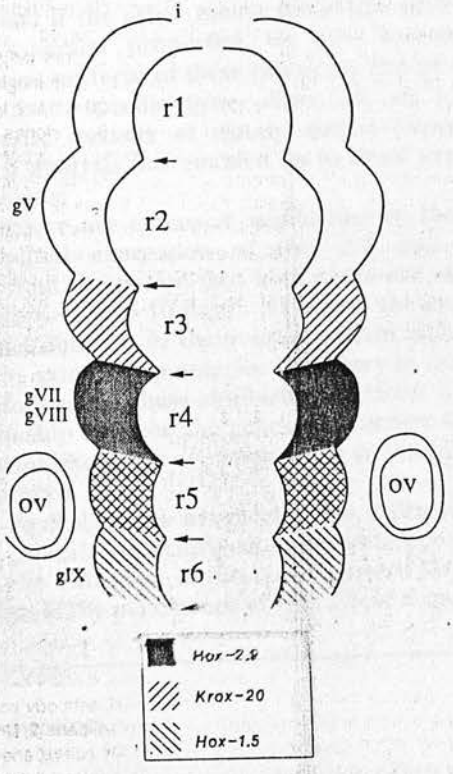


FIG. 4 Diagrammatic representation of a frontal section through the hindbrain of a 9.5-day mouse embryo showing the position of the neuromeres (rhombomeres) with respect to the otic vesicle (ov), the hindbrain–midbrain junction (isthmus, i), and the cranial ganglia which are represented alongside the rhombomeres from which they extend (gV–gIX). The arrows indicate rhombomere boundaries, and the rhombomeres are numbered from anterior to posterior (r1–r6). The shaded areas indicate the domains of expression of *Hox-2.9*, *Hox 1.5* and *Krox-20* in the hindbrain (see text for refs). *Hox-1.5* expression extends posteriorly into the neural tube.

It has been suggested that the overlapping domains of homoeobox-containing-gene expression convey positional information in the form of unique combinations of gene products in blocks along the A–P axis[4]. It remains to be seen if these blocks relate to neuromeres. In this context, *Hox-2.9* is unique in being expressed within a single neuromere, and its expression within rhombomere 4 indicates that it is involved in specifying the unique identity of this segment. □

Received 24 May; accepted 4 August 1989.

1. Akam, M. *Development* **101**, 1–22 (1987).
2. Lewis, E. B. *Nature* **276**, 565–570 (1978).
3. Kaufman, T. C. in *Time, Space and Pattern in Embryonic Development*, 365–383 (Liss, New York, 1983).
4. Holland, D. W. D. & Hogan, B. L. M. *Genes Dev.* **2**, 773–782 (1988).
5. Orr, H. *J. Morph.* **1**, 311–372 (1887).
6. Streeter, G. L. *Anat. Rec.* **2**, 111–115 (1908).
7. Neal, H. V. *J. Morph.* **31**, 293–315 (1918).
8. Adelmann, H. B. *J. comp. Neurol.* **39**, 19–171 (1925).
9. Berquist, H. & Kallen, B. *J. comp. Neurol.* **100**, 627–660 (1954).
10. Vaage, S. *Adv. Anat. Embryol. Cell Biol.* **41**, 1–88 (1969).
11. Tuckett, F., Lim, L. & Morriss-Kay, G. M. *J. Embryol. exp. Morph.* **87**, 215–228 (1985).
12. Tuckett, F. & Morriss-Kay, G. M. *J. Embryol. exp. Morph.* **88**, 231–249 (1985).
13. Sakai, Y. *Anat. Rec.* **218**, 450–457 (1987).
14. Lumsden, A. & Keynes, R. *Nature* **337**, 424–428 (1988).
15. Wilkinson, D. G., Bhatt, S., Chavier, P., Bravo, R. & Charnay, P. *Nature* **337**, 461–464 (1989).
16. Graham, A., Papalopulu, N. & Krumlauf, R. *Cell* **57**, 367–378 (1989).
17. Do, M-S. & Lonai, P. *Genomics* **3**, 195–200 (1988).
18. Deol, M. S. *J. Embryol. exp. Morph.* **12**, 475–490 (1964).
19. Hertwig, P. *Z. Indukt. Abstammungs-Vererbungsl.* **80**, 220–246 (1942).
20. Laughon, A. & Scott, M. P. *Nature* **310**, 25–31 (1984).
21. Miller, J., McLachlan, A. D. & Klug, A. *EMBO J.* **4**, 1609–1614 (1985).
22. Chavrier, P. *et al. EMBO J.* **7**, 29–35 (1988).
23. Ingham, P. *Nature* **335**, 25–34 (1988).
24. Mlodzik, M., Fjose, A. & Gehring, W. J. *EMBO J.* **7**, 2569–2578 (1988).
25. Duboule, D. & Dolle, P. *EMBO J.* **8**, 1497–1505 (1989).
26. Diederich, R. J., Merrill, V. K. L., Pultz, M. A. & Kaufmann, T. C. *Genes Dev.* **3**, 399–414 (1989).
27. Mahaffey, J. W., Diederich, R. J. & Kaufmann, T. C. *Development* **105**, 167–174 (1989).
28. Hogan, B., Holland, P. & Schofield, P. *Trends Genet.* **1**, 67–74 (1985).
29. Gaunt, S. J. *Development* **101**, 51–60 (1987).
30. Gaunt, S. J., Miller, J. R., Powell, D. J. & Duboule, D. *Nature* **324**, 662–664 (1986).
31. Davidson, D., Graham, E., Sime, C. & Hill, R. *Development* **104**, 305–316 (1988).

ACKNOWLEDGEMENTS. We thank Elizabeth Graham for help with *in situ* hybridizations, Norman Davidson for photographic expertise, Christine Sime for help in cloning *Hox-2.9*, David Porteous and Nick Hastie for critical reading of the manuscript, and Gillian M. Morriss-Kay for interpretation of early embryo sections.

CD3-negative natural killer cells express ζ TCR as part of a novel molecular complex

Paul Anderson*†, Michael Caligiuri*, Jerome Ritz* & Stuart F. Schlossman*

* Division of Tumor Immunology, Dana-Farber Cancer Institute, 44 Binney Street, Boston, Massachusetts 02115, USA
† Department of Rheumatology and Immunology, Brigham and Women's Hospital, Harvard Medical School, 75 Francis Street, Boston, Massachusetts 02115, USA

[illegible] capable of killing tumour cells in a non-MHC restricted manner[1–7], NK cells do not express cell-surface CD3, or any known target recognition structure analogous to the T cell antigen receptor (TCR) heterodimers ($\alpha\beta$ or $\gamma\delta$)[3–8]. Consistent with their lack of expression of a CD3–TCR complex, NK cells do not require prior sensitization or antigen presentation by accessory cells to specifically recognize their tumour targets[1]. Although NK cells do not express CD3–TCR, they do express CD2, the target of an alternative activation pathway which is functional in both T cells and NK cells[9–12]. In T cells, this alternative activation pathway utilizes some component of the CD3–TCR complex as a transducer molecule that is required for mitogenesis[13–15]. The fact that NK cells are activated by this alternative pathway suggested that they might express a related subunit of the CD3–TCR complex capable

Development 111, 61–74 (1991)
Printed in Great Britain © The Company of Biologists Limited 1991

Expression of the mouse *labial*-like homeobox-containing genes, *Hox 2.9* and *Hox 1.6*, during segmentation of the hindbrain

PAULA MURPHY and ROBERT E. HILL

MRC Human Genetics Unit, Western General Hospital, Edinburgh EH4 2XU, UK

Summary

The sequence of a mouse *Hox 2.9* cDNA clone is presented. The predicted homeodomain is similar to that of the *Drosophila* gene *labial* showing 80 % identity. The equivalent gene in the *Hox 1* cluster is *Hox 1.6* which shows extensive similarity to *Hox 2.9* both within and outside the homeodomain. *Hox 2.9* and *Hox 1.6* are the only two mouse members of the *labial*-like family of homeobox-containing genes as yet identified.

***Hox 2.9* has previously been shown to be expressed in a single segmental unit of the developing hindbrain (rhombomere) and has been predicted to be involved in conferring rhombomere identity. To analyse further the function of *Hox 2.9* during development and to determine if the other mouse *labial*-like gene *Hox 1.6*, displays similar properties, we have investigated the expression patterns of these two genes and an additional rhombomere-specific gene, *Krox 20*, on consecutive embryonic sections at closely staged intervals. This detailed analysis has enabled us to draw the following conclusions:**

(1) There are extensive similarities in the temporal and spatial expression of *Hox 2.9* and *Hox 1.6*, throughout the period that both genes are expressed in the embryo ($7\frac{1}{2}$ to 10 days). At 8 days the genes occupy identical domains in the neuroectoderm and mesoderm with the same sharp anterior boundary in the presumptive hindbrain. These similarities indicate a functional relationship between the genes and further suggest that the *labial*-like genes are responding to similar signals in the embryo.

(2) By 9 days the neuroectoderm expression of both genes retreats posteriorly along the anteroposterior (AP) axis. The difference at this stage between the expression patterns is the persistence of *Hox 2.9* in a specific region of the hindbrain, illustrating the capacity of *Hox 2.9* to respond to additional positional regulatory signals and indicating a unique function for this gene in the hindbrain.

(3) The restriction of *Hox 2.9* expression in the hindbrain occurs at $8\frac{1}{2}$ days, approximately the same time as *Krox 20* is first detected in the posterior adjoining domain. The mutually exclusive expression of *Hox 2.9* and *Krox 20* demarcated by sharp expression boundaries suggest that compartmentalisation of cells within the hindbrain has occurred up to 6 h before rhombomeres (morphological segments) are clearly visible.

(4) *Hox 2.9* expression is confined to the region of rhombomere 4 that shows cell lineage restriction and, unlike *Krox 20*, is expressed throughout the period that rhombomeres are visible (to $11\frac{1}{2}$ days). These data strengthen the evidence that *Hox 2.9* participates in conferring segment identity.

(5) Migrating neural crest cells that arise from rhombomere 4 are uniquely identified by the expression of *Hox 2.9* supporting the idea that neural crest cells are patterned according to their rhombomeric origin.

(6) The *Hox 1.6* gene product is differentially transcribed; only one of the two alternative transcripts codes for a homeodomain-containing protein. A comparison of the distribution patterns of the two transcripts shows that the relative proportion of homeodomain-producing message decreases as development proceeds.

Key words: homeobox, *Hox 1.6*, *Hox 2.9*, mouse development, rhombomere.

Introduction

The mouse *Antennapedia*-like homeobox genes reside within four tightly clustered multigene arrays in the mouse genome designated *Hox 1* which is on chromosome 6; *Hox 2* on chromosome 11; Hox 3 on chromosome 15; and *Hox 4* (formerly *Hox 5*, Kessel and Gruss, 1990) on chromosome 2 (Bucan *et al.* 1986; Hart *et al.* 1985; Breier *et al.* 1988; Featherstone *et al.* 1988). The clusters seem to have arisen from a common ancestral cluster by chromosomal duplication events. This is shown by the sequence comparison of the genes; for example, all but two genes in the *Hox 2* cluster have counterparts in the *Hox 1* cluster (Hart *et al.* 1987; Graham *et al.* 1989; Duboule and Dolle, 1989). The cognate genes within the clusters show a further level of

similarity in that the genes are organised in the same linear order along the chromosome. It is apparent that the mammalian clusters are remnants of an ancient ancestral cluster that predates the organisation of homeobox-containing genes in insects. The insect complex of homeobox-containing genes, the HOM-C, which includes the *Bithorax* (Lewis, 1978; Sanchez-Herrero *et al.* 1985) and *Antennapedia* (Kaufman *et al.* 1980) gene complexes of *Drosophila*, is known to be homologous to the mammalian clusters from sequence comparison and the organisation of the genes along the chromosome. At one end of the insect HOM cluster resides the gene referred to as *labial* (Diederich *et al.* 1989). At corresponding positions in the mouse *Hox 1* and *2* clusters are the homologous genes *Hox 1.6* (Baron *et al.* 1987; Mlodzik *et al.* 1988; LaRosa and Gudas, 1988) and *Hox 2.9* (Rubock *et al.* 1990); these form a subfamily of *labial*-like homeobox-containing genes. Counterparts in the other two mouse gene clusters have not been reported.

During *Drosophila* development, the embryo is divided into segments along the anteroposterior (AP) axis and *Drosophila* homeobox genes are involved in conveying positional information during the process of segmentation (Gehring, 1987 for review). In vertebrates repetitive structures in the embryonic hindbrain called rhombomeres reflect an underlying segmental organisation and act as units of cell lineage restriction (Lumsden and Keynes, 1989; Fraser *et al.* 1990). The correspondence of expression of *Hox 2.9* with a single rhombomere suggests that it plays an analogous role to *Drosophila* homeobox-containing genes in conveying positional information in the mouse (Murphy *et al.* 1989). The other members of the Hox 2 cluster occupy overlapping domains along the AP axis of the central nervous system (CNS), the anterior boundaries that lie within the hindbrain corresponding to segment boundaries (Wilkinson *et al.* 1989*b*). In general, the data available on other mouse homeobox genes show that they are expressed in the ectoderm, the mesoderm and, to a lesser extent, the endoderm of the developing embryo. Expression of the earliest of these genes is detected at the time of gastrulation (Gaunt, 1987) and a small number of genes are still expressed in the newborn mouse (Awgulewitsch *et al.* 1986; Utset *et al.* 1987). The majority of homeobox-containing genes are expressed during the process of gastrulation, when AP positional values are established in the amphibian (Ruiz i Altaba and Melton, 1990), and during morphogenesis in the embryo. Within particular developmental fields homeobox genes occupy characteristic, overlapping expression domains with different subsets of genes active in different spatial domains, consistent with a role in positional determination (Dolle *et al.* 1989; Holland and Hogan, 1988; Graham *et al.* 1989).

It appears that subfamily members in different clusters display similar, although not always identical, AP expression domains. This was observed for the domains of *Hox 3.3* (formerly *Hox 6.1*) and *Hox 1.2* (Gaunt *et al.* 1988) and for the domains of *Hox 1.4*, *Hox 2.6* and *Hox 5.1* (Gaunt *et al.* 1989) in the CNS and prevertebral column at 12.5 days. *Hox 1.5* and *Hox 2.7* both have anterior boundaries within the CNS that correspond to the same rhombomere boundary, the anterior boundary of rhombomere 5 (Gaunt *et al.* 1987; Wilkinson *et al.* 1989*a*). However, two separate studies on the related pair of genes *Hox 2.5* (Graham *et al.* 1989) and *Hox 5.2* (Duboule and Dolle, 1989) show that expression of the former extends more anteriorly. This has been interpreted as indicating that similarity between paralogues does not hold for genes expressed only in the posterior embryo (Gaunt *et al.* 1989). Some mouse homeobox-containing genes are not contained within clusters and these also have closely related genes on other chromosomes (Joyner and Martin, 1987; Davidson *et al.* unpublished data). Detailed studies of these genes, the *engrailed*-like mouse genes *En1* and *En2* (Davidson *et al.* 1988) and the *Msh*-like mouse genes *Hox 7.1* and *Hox 8.1* (Hill *et al.* 1989; Davidson *et al.* unpublished data) show that they have overlapping or complementary expression patterns.

Here we present the cDNA sequence of *Hox 2.9* and show its relationship to other *labial*-like genes. We present the results of a detailed analysis of *Hox 1.6* and *Hox 2.9* expression in the early embryo including the use of the *Krox 20* gene as a temporal and positional molecular marker in the developing hindbrain. The analysis was designed to compare the expression patterns of the *labial*-like genes in the mouse, to investigate how the segmental expression of *Hox 2.9* in rhombomere 4 becomes localised from an earlier more widespread domain and to determine how the onset of localised, segmental expression relates to the appearance of morphological segments.

Materials and methods

Isolation of cDNA clones

cDNA clones for both *Hox 1.6* and *Hox 2.9* were isolated from an 8.5 day mouse embryonic cDNA library in lambda gt10. The original clones were selected as weakly hybridising to the *Drosophila* gene *fushi tarazu*. The cDNAs were subcloned into suitable vectors for sequencing and transcription.

DNA sequence determination and analysis

Random subclones of a full-length Hox 2.9 cDNA were sequenced by dideoxynucleotide sequencing procedures (Sanger *et al.* 1987) using 'Sequenase' (US Biochemicals) as described by the manufacturers. The sequences were aligned using the Staden-plus computer package (Amersham). Sequence comparisons were carried out by the GAP program in the sequence analysis software package of the University of Wisconsin Genetics Computer Group.

Preparation of embryo sections

Embryos were obtained from outbred Swiss mice. For ageing purposes, midday on the day of detection of a vaginal plug was designated 0.5 days *post coitum*. Embryos within a litter were precisely staged by their morphology, the size and shape of the head fold and the appearance of rhombomeres being the most important criteria between 8 and 9 days. Embryos were fixed in 4% paraformaldehyde at 4°C overnight and

embedded in paraffin wax. 5–7 μm sections were cut and floated onto TESPA (3-aminopropyltriethoxysilane; Sigma) -treated slides.

In situ *hybridisation*

Sense and antisense RNA probes were produced by incorporating ^{35}S-UTP into the transcription products of selected *Hox 1.6*, *Hox 2.9* and *Krox 20* subclones inserted into T7- and T3-containing transcription vectors. The *Hox 1.6* probe used has been previously called cDNA 1 (Baron *et al.* 1987). The 3′ *Pst*I/*Eco*RI fragment of *Hox 2.9* (Fig. 1A) was used for preparations of the riboprobe. For *Krox 20* the probe was prepared from the 1.5 kb *Apa*I–*Eco*RI fragment (Chavrier *et al.* 1988). The *Hox 1.6* probe includes the homeobox sequence, but the characteristic pattern of expression observed indicates that there is no cross-reactivity under the conditions used. Sense (control) probes showed no specific labelling. The *in situ* hybridisation protocol has been previously described (Davidson *et al.* 1988) and included high-stringency washes. All probes were used at the same specific activity with $3–6\times10^5$ disints min^{-1} added per section, varied with section size.

Silver grain density estimation

The number of grains in at least two different areas of 177 μm^2 within each tissue were directly counted at times 1000 magnification. Three to six counts were taken in each area. Background counts were estimated in areas with comparable cell density and the average background count was subtracted from each count.

Results

Structure and sequence of the Hox 2.9 *gene*

A probe for the *Hox 1.6* gene, including the homeobox, was used to isolate a *Hox 2.9* cDNA from an 8.5 day mouse embryonic cDNA library by low-stringency hybridisation. We sequenced the 1780 bp *Hox 2.9* cDNA containing a single long open reading frame of 894 bp with an in-frame homeobox domain (Fig. 1). It is predicted to encode a 32×10^3 M_r protein. A *Hox 2.9* transcript of approximately 2 kb was detected in RNA prepared from 9 day embryos and F9 teratocarcinoma cells (data not shown), indicating that the sequenced cDNA clone represents an almost full-length transcript. The absence of a poly(A) stretch and recognisable polyadenylation site (Proudfoot and Brownlee, 1976) indicates that the cDNA insert is truncated at the 3′ end. The predicted ATG translational start codon is the first ATG in frame with the homeobox. No other in-frame ATG was found within 514 bp upstream of the proposed start site.

Sequence comparison reveals that *Hox 2.9* contains a homeodomain similar to that of the *Drosophila* gene *labial* (Fig. 2A). The *Drosophila labial* homeodomain has diverged significantly from the *Antennapedia* sequence (67 % amino acid identity) and is most closely related to its homologues in other species including *Hox 2.9* (Fig. 2A, Table 1). The *Hox 2.9* homeodomain shows 80 % amino acid identity to the *Drosophila labial* homeodomain and 87 % identity to that of *Hox 1.6*, which has previously been shown to be *labial*-like (85 % amino acid identity). Comparison of the *labial*-like homeodomains to other genes in the mouse shows at best 62 % identity. We therefore suggest that *Hox 2.9* is the second member of the *labial* subfamily in the mouse. No other mouse homeobox-containing gene has been reported that belongs to this subfamily; however, genes isolated from chicken (*Ghox.lab*) (Sundin *et al.* 1990) and human (*HOX 2I*) (Acampora *et al.* 1989) are *labial*-like.

Comparison of the mouse and chicken *labial*-like genes reveals that regions of similarity also exist outside the homeobox (Fig. 2B). These include a stretch of 22 amino acids at the amino-terminus of the proteins and regions that extend from both ends of the homeobox. The *Hox 1.6* gene has previously been shown to contain only two amino acids (Trp-Met) of the conserved hexapeptide (Ile/Val-Tyr-Pro-Trp-Met-Arg) found in many homeodomain proteins (Baron *et al.* 1987). However, the four replacements (Trp-Phe-Asp-X-X-Lys) in this region in *Hox 1.6* are conserved in both *Hox 2.9* and *Ghox lab* and interestingly mark the beginning of the extended region of homology around the homeodomain. Examination of the full coding region shows that *Ghox lab* is more similar to *Hox 2.9* (Fig. 2B, Table 1), particularly at the carboxy-terminal end, and we predict that this chicken gene is the *Hox 2.9* homolog; however, expression analysis is required to make definite conclusions.

Hox 2.9 *and* Hox 1.6 *expression between* $7\frac{1}{2}$ *and 9 days of development (formation of rhombomeres)*

Between $7\frac{1}{2}$ and 9 days of development, the expression patterns of *Hox 2.9* and *Hox 1.6* change rapidly and dramatically. To establish how the patterns evolve during this period, we have examined embryos at closely staged intervals of approximately 6 h, using the *Krox 20* (Wilkinson *et al.* 1989*a*) gene as a molecular marker for events occurring in the developing hindbrain. Expression of *Hox 2.9* and *Hox 1.6* is first detected at $7\frac{1}{2}$ days during the early stages of gastrulation (Fig. 3). Both genes are expressed within the primitive streak in newly formed mesoderm and overlying neuroectoderm. *Hox 2.9* expression is at a higher level and is more extensive than *Hox 1.6*. In the early 8 day embryo, *Hox 2.9* and *Hox 1.6* are expressed at a high level in extensive domains, extending from the posterior end of the embryo along the neuroectoderm and mesoderm tissue layers into the region of the developing hindbrain of the headfold (Fig. 4A–D). The two genes have identical, sharp anterior boundaries of expression in the neuroectoderm that coincide with the preotic sulcus (a characteristic groove in the surface of the presumptive hindbrain). An adjacent section probed with the *Krox-20* gene shows that it is expressed in a single domain in the hindbrain, the posterior boundary of which corresponds to the anterior boundary of *Hox 2.9* and *Hox 1.6* (Fig. 4D). As was previously described, *Krox 20* is first detected in a single domain in the hindbrain and later in its characteristic two-stripe pattern (Wilkinson *et al.* 1989*a*). Within the mesoderm of the 8-day embryo, both *Hox 2.9* and *Hox 1.6* expression is restricted to lateral plate mesoderm as

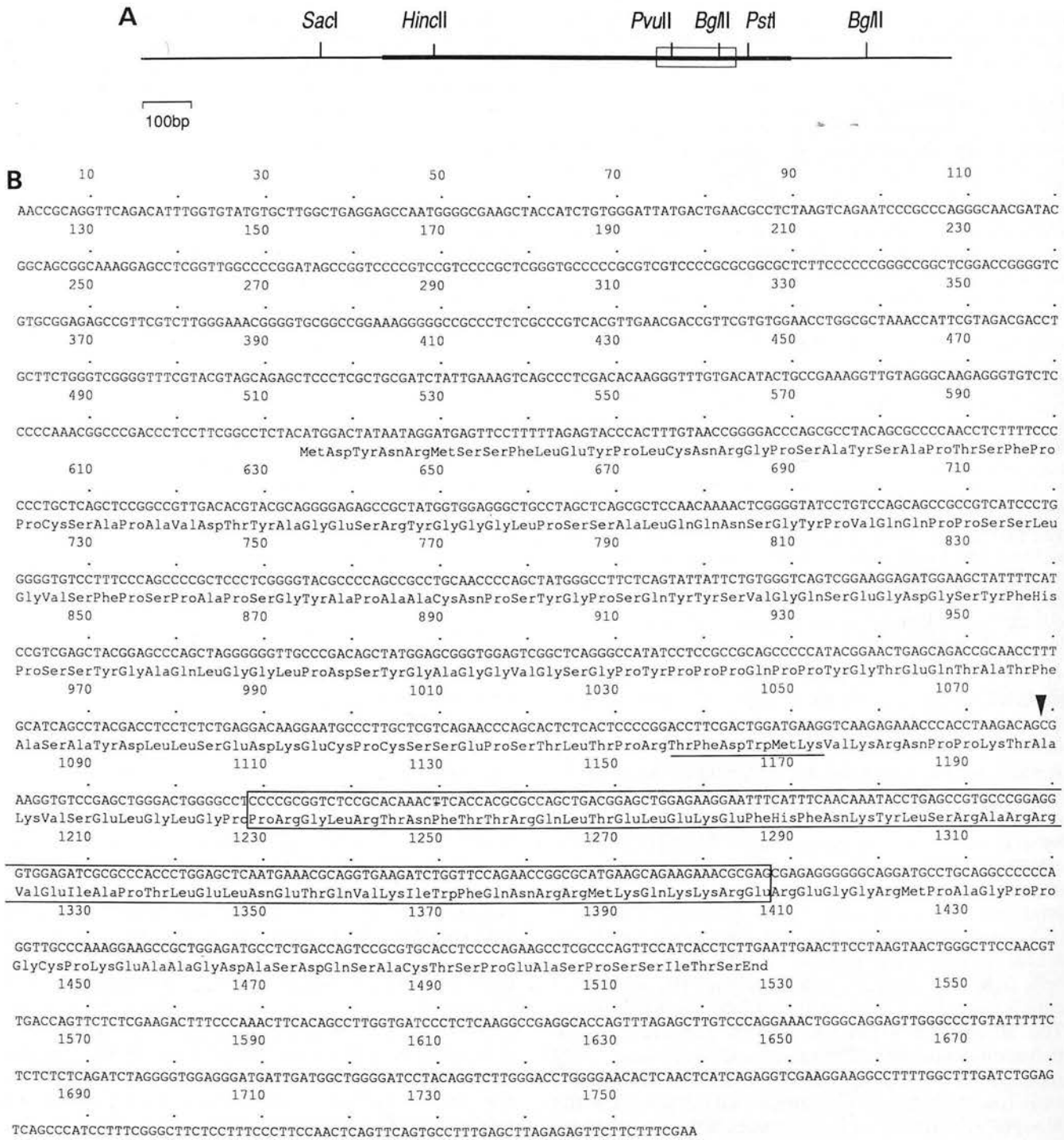


Fig. 1. The *Hox 2.9* cDNA clone. (A) A line diagram of the 1759 bp *Hox 2.9* cDNA. The heavy line represents the predicted coding sequence. The open box represents the homeobox. Key restriction enzyme sites are indicated. (B) The sequence of the *Hox 2.9* cDNA illustrated in A. The amino acid sequence of the longest open reading frame with the homeodomain in frame (boxed) is given below the DNA sequence. The conserved hexapeptide is underlined. The arrowhead shows the predicted splice site based on comparison with other *labial*-like gene sequences.

far anterior as the headfold (Fig. 5D–F) and to presomitic mesoderm with expression decreasing as somites condense.

By $8\frac{1}{2}$ days of development, *Hox 2.9* expression in the hindbrain has become localised (Fig. 4G). The anterior boundary is at the same position as at 8 days (the preotic sulcus) with *Hox 2.9* and *Krox 20* continuing to share this boundary, but now there is a new posterior boundary also within the hindbrain. We simultaneously detect the initiation of the second band of *Krox 20* expression, the anterior boundary of which coincides with the posterior boundary of *Hox 2.9*. The expression

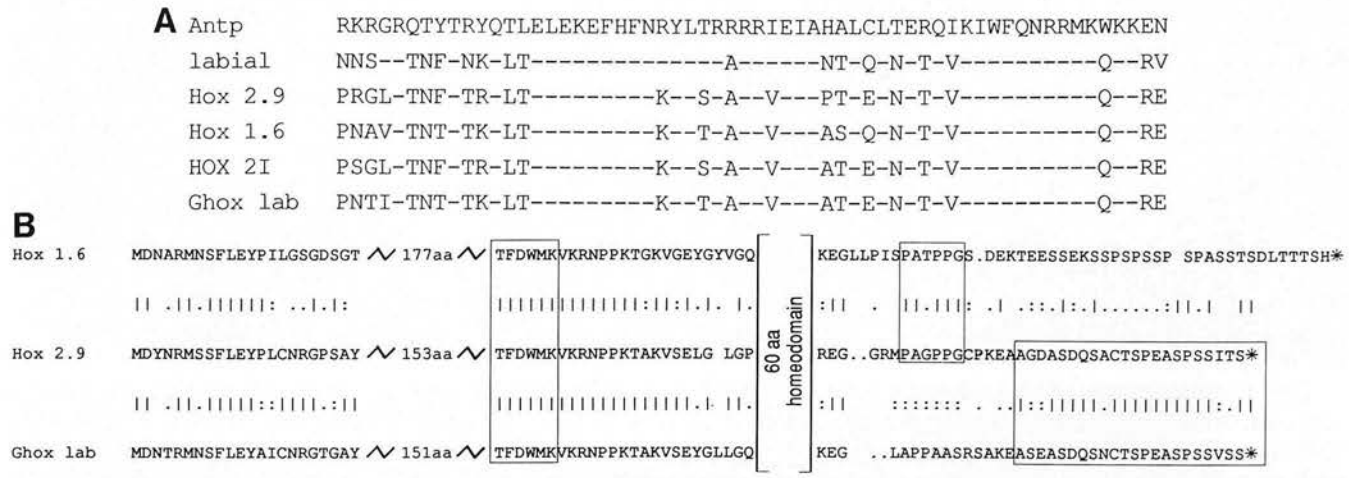


Fig. 2. Alignment of the amino acid sequences of *labial*-like genes. (A) The amino acid sequences of the homeodomains of *Drosophila labial* and the vertebrate *labial*-like genes (*Hox 2.9* and *Hox 1.6*: mouse, *HOX 2I*: human, *Ghox lab*: chicken, see text for references) compared to *Drosophila Antennapedia*. Differences from the *Antennapedia* sequence are noted revealing the characteristic features of the *labial* family of genes. (B) A representation of the full-coding region of *Hox 1.6*, *Hox 2.9* and *Ghox lab*. A line between two sequences represents an amino acid identity. Dots represent conservative changes (as judged by the GAP program). The conserved hexapeptide is boxed. A divergent stretch 5′ of the hexapeptide, and the homeodomain shown in 2A, are omitted from this diagram.

Table 1. *Comparison of* labial *and* labial-*like genes*

	HOX 2I		*Hox 1.6*		*Ghox lab*		*labial*	
(a) *Homeobox/homeodomain comparisons*								
	DNA	amino acid	DNA	amino acid	DNA	amino acid	DNA	amino acid
Hox 2.9	87%	96.6%	79.0%	86.6%	80.6%	90.0%	72.3%	80.0%
HOX 2I			82.8%	88.3%	82.8%	91.7%	74.5%	81.6%
Hox 1.6					81.1%	93.3%	76.2%	85.0%
Ghox lab							74.5%	85.0%
(b) *Whole protein comparisons*								
	Similarity	Identity	Similarity	Identity	Similarity	Identity		
Hox 2.9	90.3%	85.6%	62.5%	45%	71.1%	55.1%		
Hox 1.6					58.8%	46.5%		

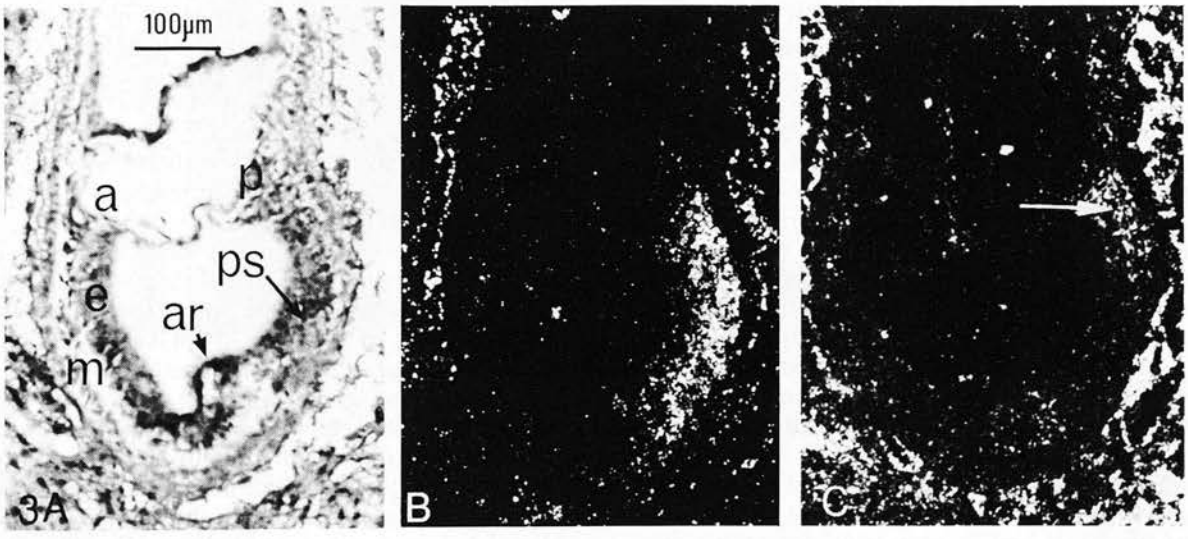


Fig. 3. Adjacent sagittal sections through a 7½ day mouse embryo. A and B show bright- and dark-field images of the same section probed with *Hox 2.9*. C was probed with *Hox 1.6*. a, anterior; p, posterior; ps, primitive streak; e, ectoderm; m, mesoderm; ar, archenteron. The arrows indicate the labelled cells in C and D.

of *Hox 2.9* in the anterior neural tube seems to retreat posteriorly at this time with expression persisting in more posterior regions. *Hox 1.6* expression also appears to retreat posteriorly along the neural tube in the same way; however, in contrast to *Hox 2.9* no expression of *Hox 1.6* remains in the hindbrain (Fig. 4F). Within the mesoderm both *Hox 2.9* and *Hox 1.6* remain expressed in lateral plate mesoderm up to the level of the posterior hindbrain and in presomitic mesoderm. We now first detect expression of both genes in an endodermal derivative, the epithelium of the foregut pocket.

The rhombomeres, which are the morphological representation of segments within the hindbrain, are visible at $8\frac{3}{4}$ days (Fig. 4I). The rhombomeres are small and more evenly shaped at this stage than at later stages. We can now see that the expression domains of *Hox 2.9* and *Krox 20* within the developing hindbrain are perfectly coincident with rhombomere 4 in the case of *Hox 2.9* and rhombomeres 3 and 5 in the case of *Krox 20*. These results show that the expression of *Hox 2.9* and *Krox 20* becomes localised within the hindbrain in an anterior-to-posterior order up to 6 h before segments are visible.

At $8\frac{3}{4}$ days, labelling with *Hox 2.9* is also detected in the mesoderm lateral to rhombomere 4 in the region where sensory ganglia are condensing (Fig. 4J). Migrating neural crest cells that originate from rhombomere 4 also express *Hox 2.9* (Fig. 5A,B).

Hox 2.9 *and* Hox 1.6 *expression between 9 and 11 days*

Between 9 and 10 days of development *Hox 2.9* and *Hox 1.6* are expressed within the neural tube in posterior regions only, in a way that is consistent with the expression domains retreating posteriorly, since there is a posterior-to-anterior gradient (Fig. 5J–L). This may relate to the process of maturation in the neural tube. A dorsoventral gradient of *Hox 2.9* expression within the neural tube is also visible (not shown) and this relates to a period of cytodifferentiation in which sensory neurons are being produced in the dorsal region of the neural tube where *Hox 2.9* is most abundantly expressed. Dorsoventral sublocalisation of homeobox gene expression within the neural tube has previously been described (Bogarad *et al.* 1989). *Hox 2.9* is expressed most heavily within rhombomere 4 of the hindbrain (Fig. 5G,H). We have previously described how sharply defined this domain is at the cellular level (Murphy *et al.* 1989). A series of sections through a 10-day embryo shows that a very narrow band of cells in the floor plate of rhombomere 4 does not express *Hox 2.9* (Fig. 6), this complements the fact that rhombomere boundaries do not extend into the floor plate.

Within the mesoderm, expression of both genes is now seen in gut-associated mesoderm at and below the level of the heart and in remaining presomitic mesoderm in posterior regions (Fig. 5G–L). In addition, *Hox 2.9* is expressed in the nephrogenic duct of the developing kidney (Fig. 5K). The domains of the two genes in the gut-associated mesoderm have the same AP restrictions, although *Hox 1.6* expression appears to be more extensive laterally, but this may simply reflect differences in the efficiencies of the two probes. There is also expression in the surface ectoderm adjacent to the labelled gut-associated mesoderm (Fig. 5J–L). Both genes are expressed in gut epithelium at the level of the forelimb bud (Fig. 5G–I). This is a derivative of the endoderm and is therefore one of the few examples of endodermal expression of homeobox-containing genes (Holland and Hogan, 1988; Duprey *et al.* 1988).

By $10\frac{1}{2}$ days the mesodermal expression of *Hox 2.9* and *Hox 1.6* has been down-regulated (Fig. 7A–C) and is not detectable at $11\frac{1}{2}$ days (Fig. 7D–F). At $11\frac{1}{2}$ days, the rhombomeres are no longer visible and the expression of *Krox 20* is no longer detectable but *Hox 2.9* expression persists at a reduced level in the hindbrain (Fig. 7E). By $12\frac{1}{2}$ days no expression of *Hox 2.9* or *Hox 1.6* is detectable in the embryo.

Expression of the differential transcripts of Hox 1.6

Hox 1.6 is differentially spliced to give two transcripts that differ by a 203 bp region 5′ of the homeobox (LaRosa and Gudas, 1988). The transcripts that contain this region code for a full-length protein with a homeodomain whereas the transcripts that lack this region code for a truncated protein with the same amino-terminal half but no homeodomain. In F9 teratocarcinoma cells both transcripts are produced, the relative amount of the shorter transcript increasing from 10 % to as high as 56 % after treatment with retinoic acid; a treatment that induces the cells to differentiate (LaRosa and Gudas, 1988). We isolated 17 different *Hox 1.6* cDNA clones from an $8\frac{1}{2}$ day cDNA library and found that only 10 contained the differentially spliced region, showing that both forms of transcript are produced in the early embryo. Following specific amplification by polymerase chain reaction, the differentially spliced region was subcloned into a transcription vector to produce antisense and sense RNA for *in situ* hybridisation. This probe is referred to as the differential probe *Hox 1.6.d* and hybridises to only full-length transcripts encoding the homeodomain. The expression pattern detected by *Hox 1.6d* was compared to that observed with a 3′ probe, which hybridises to both transcripts, by analysis of consecutive embryo sections.

At 8 days of development both *Hox 1.6* probes detected the same widespread domain of expression (Fig. 8A–C). The labelling with *Hox 1.6.d* was at a lower level (53–65 %, Table 2) but it shows that the full-length transcript is being produced in the embryo at 8 days. At 9 days of development, however, when the *Hox 1.6* 3′ probe detects transcripts in a broad region of gut-associated mesoderm and gut epithelium, presomitic mesoderm, and posterior neural tube, the full-length transcript is only detectable above background with *Hox 1.6.d* in the gut epithelium (Fig. 8D–G). The labelling of the gut epithelium with *Hox 1.6d* is too low to be visible in the photographs, but the grain counts show that it is labelled above background. Estimates

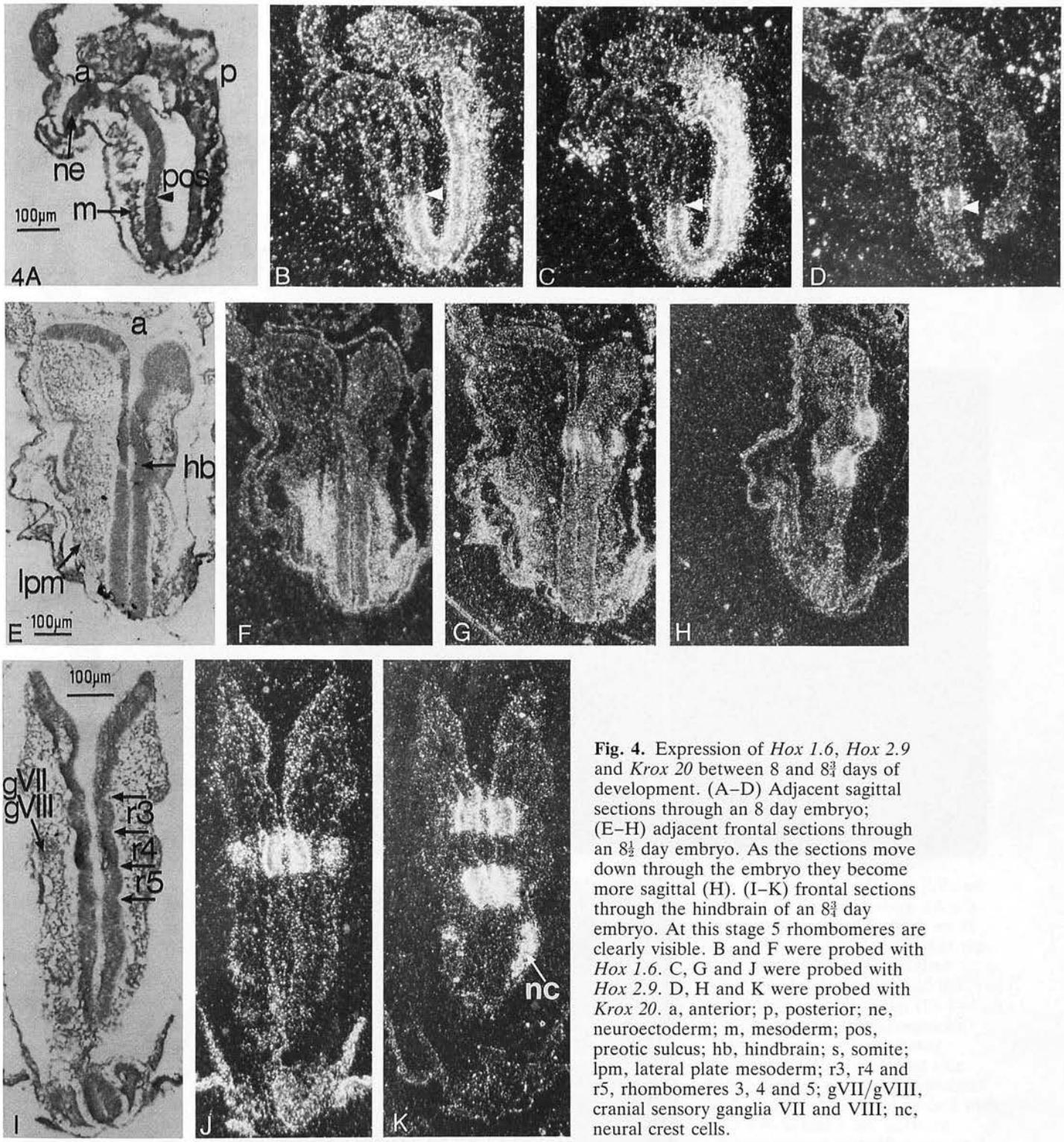


Fig. 4. Expression of *Hox 1.6*, *Hox 2.9* and *Krox 20* between 8 and $8\frac{3}{4}$ days of development. (A–D) Adjacent sagittal sections through an 8 day embryo; (E–H) adjacent frontal sections through an $8\frac{1}{2}$ day embryo. As the sections move down through the embryo they become more sagittal (H). (I–K) frontal sections through the hindbrain of an $8\frac{3}{4}$ day embryo. At this stage 5 rhombomeres are clearly visible. B and F were probed with *Hox 1.6*. C, G and J were probed with *Hox 2.9*. D, H and K were probed with *Krox 20*. a, anterior; p, posterior; ne, neuroectoderm; m, mesoderm; pos, preotic sulcus; hb, hindbrain; s, somite; lpm, lateral plate mesoderm; r3, r4 and r5, rhombomeres 3, 4 and 5; gVII/gVIII, cranial sensory ganglia VII and VIII; nc, neural crest cells.

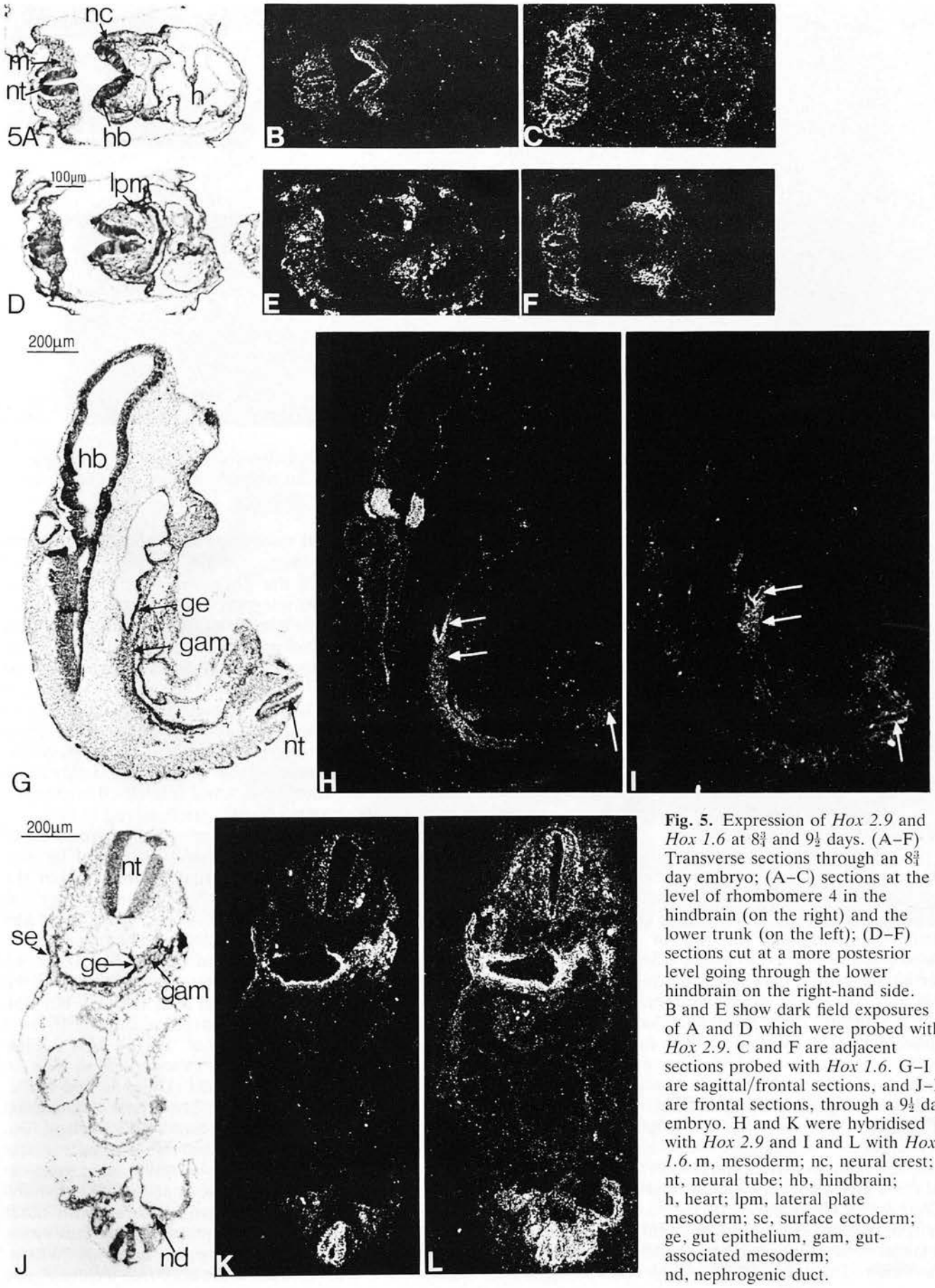


Fig. 5. Expression of *Hox 2.9* and *Hox 1.6* at $8\frac{3}{4}$ and $9\frac{1}{2}$ days. (A–F) Transverse sections through an $8\frac{3}{4}$ day embryo; (A–C) sections at the level of rhombomere 4 in the hindbrain (on the right) and the lower trunk (on the left); (D–F) sections cut at a more postesrior level going through the lower hindbrain on the right-hand side. B and E show dark field exposures of A and D which were probed with *Hox 2.9*. C and F are adjacent sections probed with *Hox 1.6*. G–I are sagittal/frontal sections, and J–L are frontal sections, through a $9\frac{1}{2}$ day embryo. H and K were hybridised with *Hox 2.9* and I and L with *Hox 1.6*. m, mesoderm; nc, neural crest; nt, neural tube; hb, hindbrain; h, heart; lpm, lateral plate mesoderm; se, surface ectoderm; ge, gut epithelium, gam, gut-associated mesoderm; nd, nephrogenic duct.

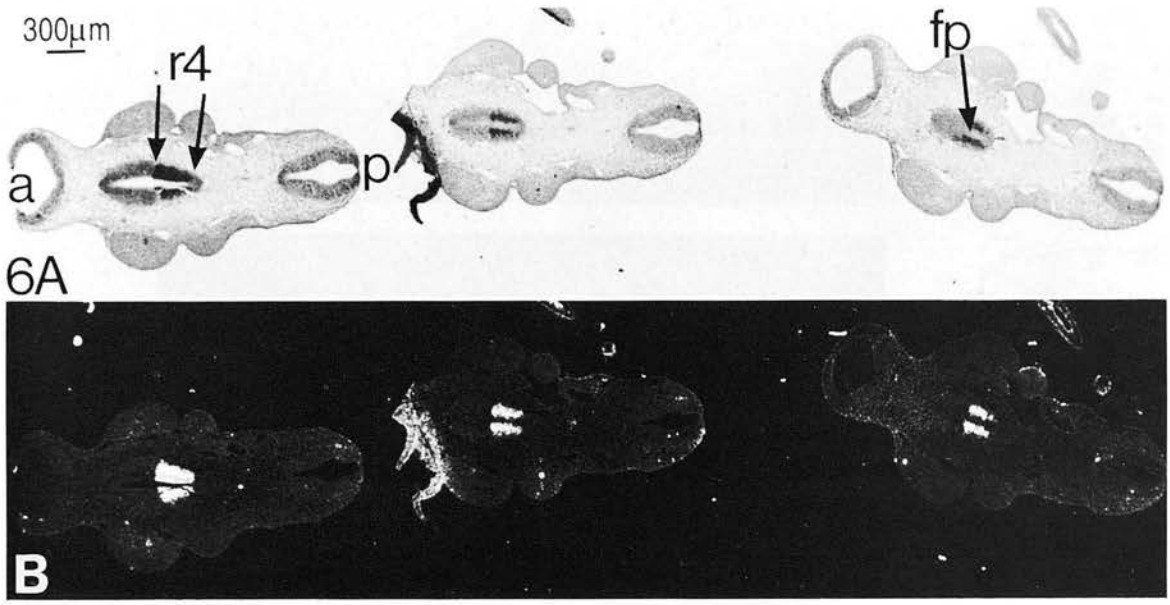


Fig. 6. *Hox 2.9* expression in a series of consecutive frontal sections through a 10½ day embryo. The sections to the right are progressively more ventral and include the floor plate of the hindbrain which is not labelled with *Hox 2.9*. a, anterior; p, posterior; r4, rhombomere 4; fp, floorplate.

from direct silver grain counts show that there is an overall drop in the proportion of full-length transcripts produced at 9 days. The three areas that were examined (Table 2); gut-associated mesoderm, gut epithelium and posterior neural tube, all show dramatic decreases in the level of labelling. The counts for gut-associated mesoderm and posterior neural tube were not above background counts.

Discussion

Evolution of the labial *family of genes*

Hox 2.9 and *Hox 1.6* (Mlodzik *et al.* 1988; LaRosa and Gudas, 1988) are the mouse homologs of the *Drosophila* gene *labial*. *labial*-like genes have also been identified in the human and the chicken. The homeodomains of the vertebrate genes are very similar to that of *labial* (80–85 % identical, Table 1), but there is little similarity throughout the rest of the protein. The *Drosophila* protein is 629 amino acids long whereas the vertebrate proteins are much shorter varying between 298 and 336 amino acids. Vertebrates also lack the intron that interrupts the homeodomain of the *Drosophila* gene (Diederich *et al.* 1989; LaRosa and Gudas, 1988; Acampora *et al.* 1989). The vertebrate genes are more similar to each other in structure and sequence with homology extending outside the homeodomain (Fig. 2B). Comparing the full-length proteins, it appears that the chicken gene *Ghox-lab* is more similar to *Hox 2.9* than *Hox 1.6* although the homeodomains of all three genes are very similar. It seems therefore that *Ghox-lab* is the homolog of *Hox 2.9*. The expression pattern of *Ghox-lab* has not been fully described and it remains to be seen if it is segmentally expressed in the hindbrain. This information will be valuable since the chicken is a useful system for developmental manipulation.

The genetics of the *Drosophila* gene *labial* have proven difficult to interpret but a homeotic role for *labial* has been concluded from clonal studies (Merrill *et al.* 1989). The *labial* protein has been found in neural and epidermal cells of a very distinct region of the head that is thought to represent an ancestral segment (Diederich *et al.* 1989). The fact that *labial* and one of its mouse homologs, *Hox 2.9*, are expressed in single anterior segments is striking. Although it is likely that there are differences in the systems for determining position in two such distinct and specialised organisms, these highly conserved genes are involved in both.

An interesting general feature of the expression of clustered homeobox genes, which is shared by vertebrates and *Drosophila*, is that position within the cluster is reflected in position along the body axis at which the gene is expressed (Akam, 1987; Scott and Carroll, 1987; Harding *et al.* 1985; Graham *et al.* 1989; Duboule and Dolle, 1989). In this respect, *Hox 2.9* represents a special case in that it is positioned at the end, termed the 3′ end, of the cluster, but the neighbouring gene to the 5′ side, *Hox 2.8*, is expressed more anteriorly (Wilkinson *et al.* 1989*b*). *Hox 2.8* has no equivalent gene in the *Hox 1* cluster and so *Hox 1.6* is the most anteriorly expressed (Duboule and Dolle, 1989). Both mouse *labial*-like genes have exceptional expression patterns within the hindbrain at 9 days. The expression of *Hox 1.6* in the hindbrain is more transient than that of the other homeobox-containing genes in that no expression is detectable at 9 days. *Hox 2.9* expression disrupts the pattern observed with other *Hox 2* cluster members of sequential genes possessing anterior boundaries at two-segment intervals (Wilkinson *et al.* 1989*b*). *Hox 2.9* is the only *Hox 2* cluster gene

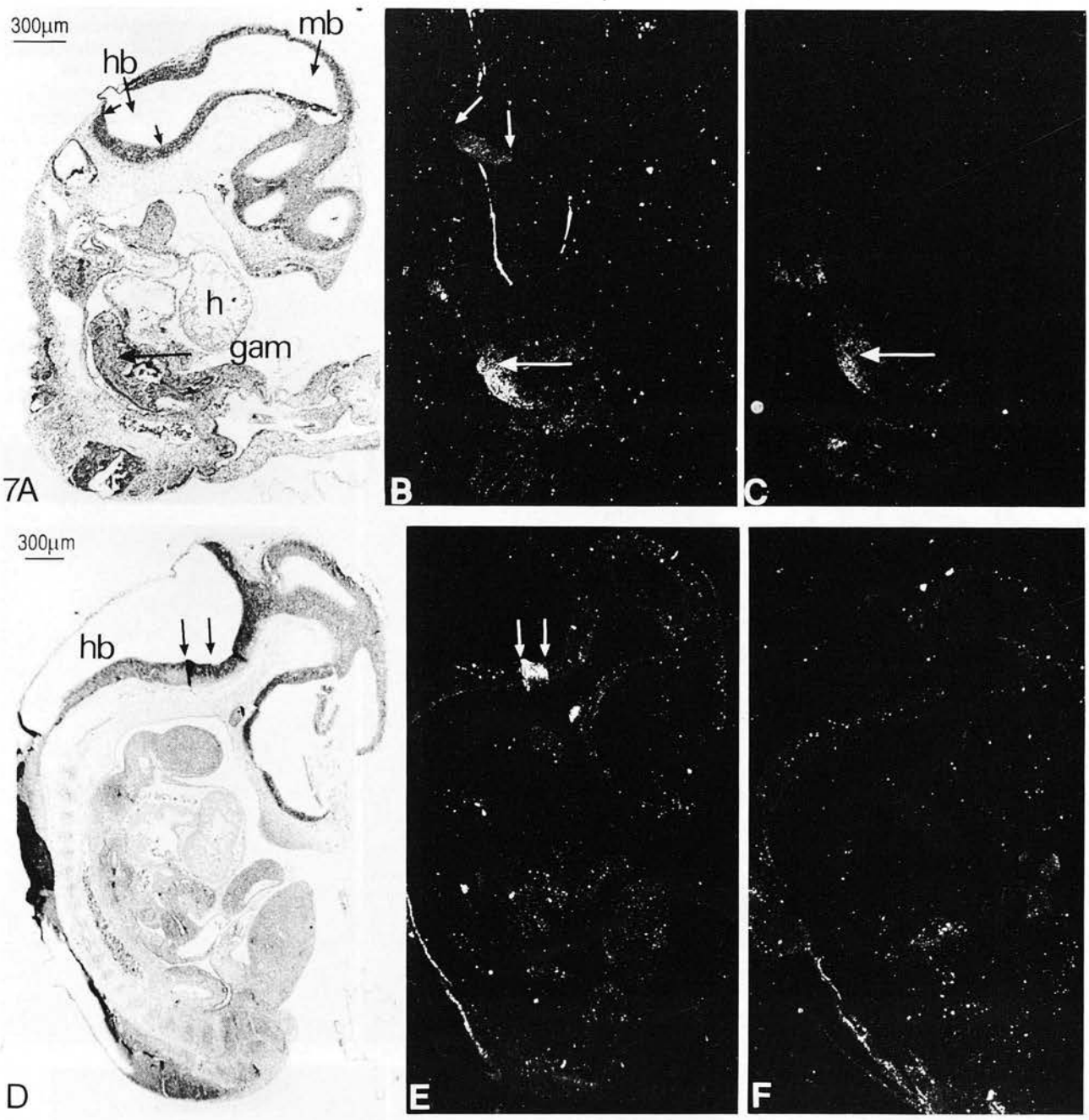
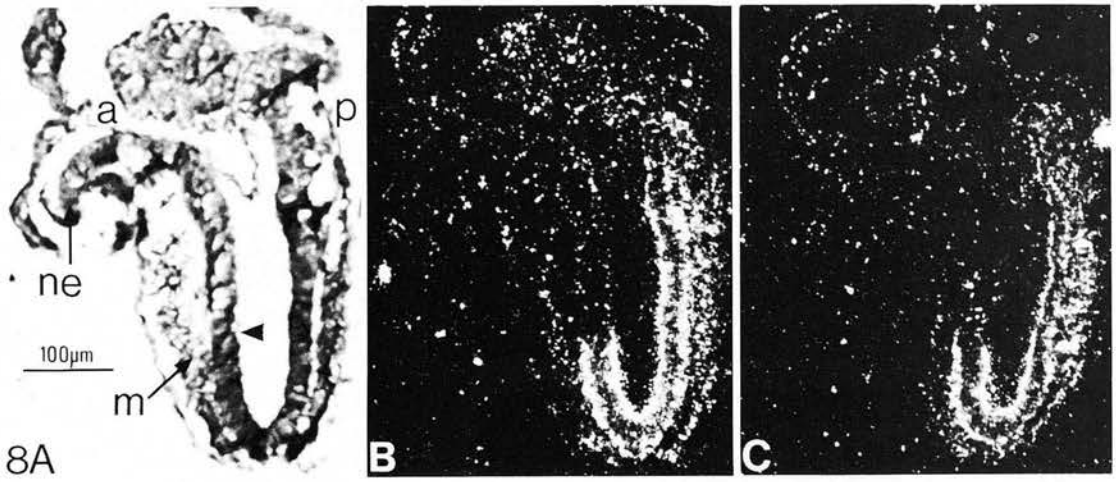
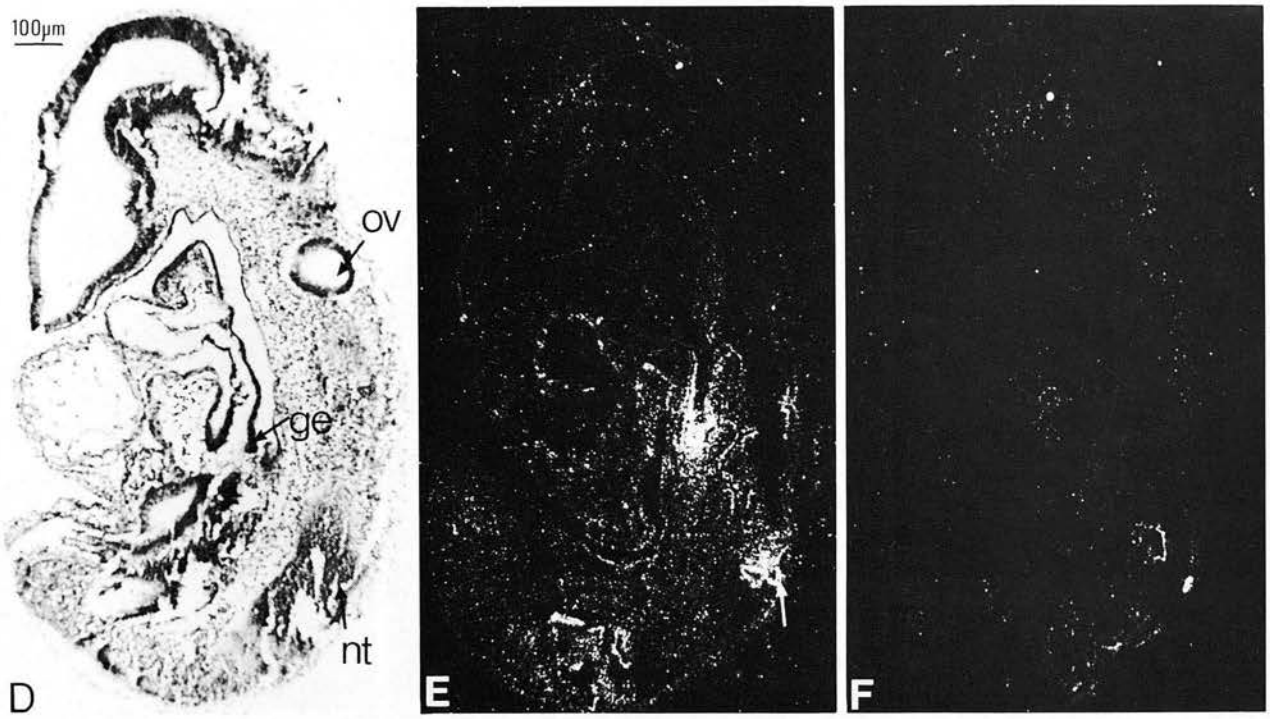


Fig. 7. Expression of *Hox 2.9* and *Hox 1.6* between $10\frac{1}{2}$ and $11\frac{1}{2}$ days of development. A–C show sagittal sections through a $10\frac{1}{2}$ day embryo. D–F show sagittal sections through an $11\frac{1}{2}$ day embryo. B and E show labelling with *Hox 2.9*; C and F have been probed with *Hox 1.6*. a, anterior; p, posterior; hb, hindbrain; mb, midbrain; h, heart; gam, gut-associated mesoderm.

to be uniquely expressed in a single rhombomere. Unlike other *Hox* genes which are generally expressed in overlapping domains in the somites and later the prevertebral column (Holland and Hogan, 1988; Graham *et al.* 1989; Duboule and Dolle, 1989), mouse *labial*-like genes are only expressed in the presomitic mesoderm with expression decreasing as somites condense.

As a result of duplication there are at least two *labial*-like genes in the mouse, and most likely also in other vertebrates. The duplicated genes have diverged but have maintained to a remarkably high degree certain features in common; this in itself would indicate a conservation of function at some level. This study has also shown that the two mouse genes share several features of their expression patterns, which extends the theory of functional similarity. The duplication of these genes must have been necessary to accommodate the


a
p
ne
100μm
m
8A
B
C


100μm
ov
ge
nt
D
E
F

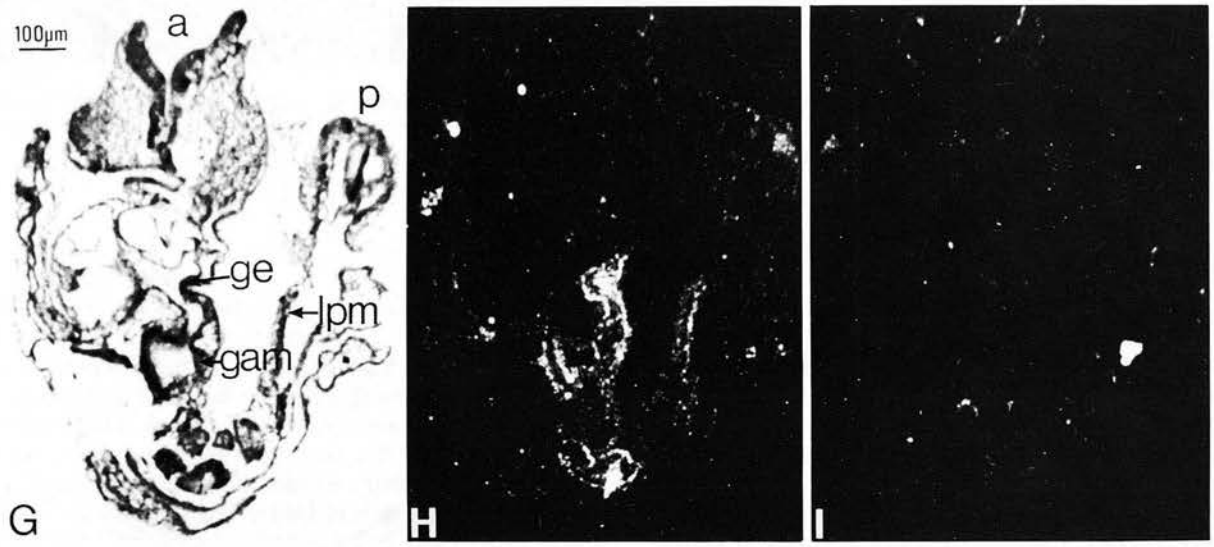

100μm
a
p
ge
lpm
gam
G
H
I

Fig. 8. Expression of the differential transcripts of *Hox 1.6* at 8 and 9 days of development. (A–C) Sagittal sections through an 8 day embryo. (D–F) Sagittal sections, and (G–I) frontal sections through a 9 day embryo. B, E and H were probed with a fragment from the 3′ end of *Hox 1.6* which detects both transcripts. C, F and I were probed with *Hox 1.6d* which detects only the full-length transcript. The arrowhead denotes the anterior boundary of expression; a, anterior; p, posterior; ne, neuralectoderm; m, mesoderm; ov, otic vesicle; ge, gut epithelium; nt, neural tube; gam, gut-associated mesoderm; lpm, lateral plate mesoderm.

establishment of the more complex body plan of the vertebrate in which members of subfamilies maintain similar although specialised functions.

Hox 2.9 *and* Hox 1.6 *expression: temporal and spatial similarities*

Hox 2.9 and *Hox 1.6* share several features of their temporal and spatial expression patterns. Expression of both genes is initiated at the posterior end of the $7\frac{1}{2}$ day gastrulating embryo. By 8 days, the expression patterns are indistinguishable with both genes occupying domains that extend from the primitive streak into the presumptive hindbrain, with identical anterior boundaries within the neuroectoderm at the preotic sulcus. At a relatively early stage (before $8\frac{1}{2}$ days), the neural tube expression retreats along the AP axis so that at $9\frac{1}{2}$ days only very posterior regions of the neural tube express *Hox 2.9* and *Hox 1.6*, with the additional persistent expression of *Hox 2.9* in a distinct region of the hindbrain. There are also parallels in the mesodermal and ectodermal expression of *Hox 2.9* and *Hox 1.6*; in the presomitic and lateral plate mesoderm at $8\frac{1}{2}$ days where transverse sections reveal that they have the same AP limits; in the modified mesodermal expression at $9\frac{1}{2}$ days when both genes are expressed in gut-associated mesoderm; and in the lack of detectable mesodermal expression after 10 days. The timing, the extent and the transient nature of the expression of these two genes indicate that they are responding to the same or similar signals in the embryo. The rapid loss of *Hox 2.9* and *Hox 1.6* RNA from the neural tube is either due to a loss of this stimulating signal or to simultaneous active repression of the genes. It also indicates that the transcripts have a rapid turnover rate, which has been shown for *Drosophila ftz* RNA with a half-life of 6–8 min (Edgar *et al.* 1986). The persistence of *Hox 2.9* expression in the hindbrain indicates that it can respond to an additional specific signal related to the segmentation of the hindbrain. For this reason, it will be important to investigate the control regions of these genes and to compare the binding sites for regulators that are present.

During the early phase of *labial*-like gene expression ($7\frac{1}{2}$–$8\frac{1}{2}$ days), there appears to be coordinate expression in the ectodermal and mesodermal tissue layers resulting in corresponding AP limits in these two tissues. As development proceeds and the complexity of the embryo increases, expression in both tissue layers becomes modified and there is little correspondence between the two layers. We therefore suggest that in the early stages of development basic AP positional domains are being similarly defined in the embryo as a whole whereas in the later embryo developmental fields become more independent.

The unique expression of Hox 2.9 within rhombomere 4 must represent a specialised function for *Hox 2.9*. It is possible that the earlier more widespread domain, from which this domain is derived and which is shared with *Hox 1.6*, is only functional in priming the later restricted expression. This is reminiscent of *Drosophila* homeotic genes, which have early widespread expression domains that become restricted to the corresponding functional domains (for review Akam, 1987). Alternatively the broad expression of *Hox 2.9* and *Hox 1.6* at 8 days may be involved in positional signalling that is important prior to segmentation.

The relationship between the expression of Hox 2.9 *and hindbrain segmentation*

The localised expression of *Hox 2.9* and *Krox 20* within specific rhombomeres has previously been described (Murphy *et al.* 1989; Wilkinson *et al.* 1989*a*). The analysis presented here focuses on earlier *Hox 2.9* expression and details further the role that this gene plays in hindbrain segmentation. Furthermore, together with *Krox 20*, these genes provide useful molecular markers in studying the process of segmentation of the hindbrain. *Krox 20* is first expressed in two domains within the hindbrain that will become rhombo-

Table 2. *Comparison of grain counts from embryo sections representing differential transcripts of* Hox 1.6

	Hox 1.63′ probe*	Hox 1.6d†	Hox 1.6d/Hox 1.63′
8 day neuroectoderm	91.3±2.9	59.1±8.4	0.65
8 day mesoderm	84.1±5.1	44.6±4.3	0.53
9 day gut epithelium	130.0±6.9	26.1±12.2	0.20
9 day neural tube	60.3±4.9	4.5±2.7‡	0.08
9 day mesoderm	40.2±10.6	4.3±4.6‡	0.10

Silver grain counts from *in situ* hybridised embryo sections (see Materials and methods) and the ratio of mean counts with two Hox 1.6 probes.

*Hox 1.63′ probe detects both differential transcripts of Hox 1.6.

†Hox 1.6d only detects full length transcripts. The differences between the mean counts for the two probes were found to be statistically signficant at less than the 1% level in all cases. With the exception of the values marked (‡) all were significantly above background estimations.

meres 3 and 5. This expression is first initiated in the more anterior domain, followed by initiation in the more posterior domain with both domains expressing *Krox 20* prior to the appearance of rhombomeres (Wilkinson *et al.* 1989*b*). The results presented here (Fig. 9) show that *Hox 2.9* is expressed in the hindbrain, with a defined anterior boundary, at the time that *Krox 20* is expressed in a single domain and before rhombomeres are visible. This is in contrast to the findings of Wilkinson *et al.* (1989*b*) who suggested that *Hox 2.9* is not expressed in the hindbrain when *Krox 20* is first detected. We further find that *Hox 2.9* expression becomes localised to the region of the hindbrain that will form rhombomere 4 at approximately the same time that *Krox 20* expression is initiated in the second domain.

At no time did we observe an overlap in the expression of *Hox 2.9* and *Krox 20* and once the domains are established they have sharp planar boundaries indicating that there is little or no cell mixing occurring between the domains. The expression pattern of these genes would therefore indicate that compartmentalisation of the hindbrain begins in the 8 day embryo and progresses in an anterior-to-posterior direction. By $8\frac{1}{2}$ days the segmental units represented by rhombomeres 3, 4 and 5 have been defined. After rhombomeres are visible, we show that in rhombomere 4 *Hox 2.9* is not expressed in the floorplate. Fraser *et al.* (1990) have demonstrated that there are no rostrocaudal cell lineage restrictions in the floor plate of the chick hindbrain and that the floor plate also lacks visible rhombomere boundaries. *Hox 2.9* is therefore only expressed in the part of rhombomere 4 that is obviously segmented. Rhombomeres are transient structures that disappear by day 12. The *Hox 2.9* rhombomere 4 expression is not detectable after $11\frac{1}{2}$ days and therefore expression persists throughout the period that rhombomere 4 exists. These data further suggest that *Hox 2.9* is involved in specifying the identity of the developmental compartment defined as rhombomere 4.

Hox 2.9 is also expressed in the sensory ganglia associated with rhombomere 4 and in the neural crest cells that migrate from rhombomere 4. The expression of *Hox 2.9* specifically in the neural crest cells that arise from rhombomere 4 supports the idea of neural crest cells being patterned according to their rhombomeric origin (Couly and LeDouarin, 1990). Neural crest cells are known to follow specific migratory pathways maintaining the AP order in which they arise (Tan and Morriss-Kay, 1986; Noden, 1975). Patterning of neural crest cells according to rhombomeric origin would therefore extend the segmental unit to regions outside the neuroectoderm.

It appears that the development of the hindbrain is under a complex regime of regulatory controls since the expression patterns of *Hox 2.9*, *Hox 1.6* and *Krox 20* within the developing hindbrain follow different modes as segmentation occurs. Both *Hox 2.9* and *Hox 1.6* are expressed early in broad domains with sharp anterior boundaries within the hindbrain; *Hox 2.9* later becomes localised to a single segment at which time *Hox 1.6* is no longer expressed. *Hox 2.9* and *Krox 20* are expressed in a segmental pattern; however, *Hox 2.9* results from the modification of a broad region of expression and *Krox 20* is initiated in distinct domains. Whereas the segmental expression of *Hox 2.9* and *Krox 20* appears to be established at the same time, *Hox 2.9* expression persists for a longer period. These genes will be important in understanding the positional signalling events and the regulatory elements involved in the process of hindbrain segmentation.

Hox 1.6 *differential transcripts; a change in their relative proportions as development proceeds*

The alternate *Hox 1.6* transcripts produced in the embryo code for a homeodomain protein and a truncated non-homeodomain protein. It is possible that the truncated protein has an independent function in the developing embryo or production of the spliced RNA may simply be a means of silencing the *Hox 1.6* direct DNA-binding function. It is not possible at present to tell if the truncated protein is functional and if so what that function might be. Although it does not

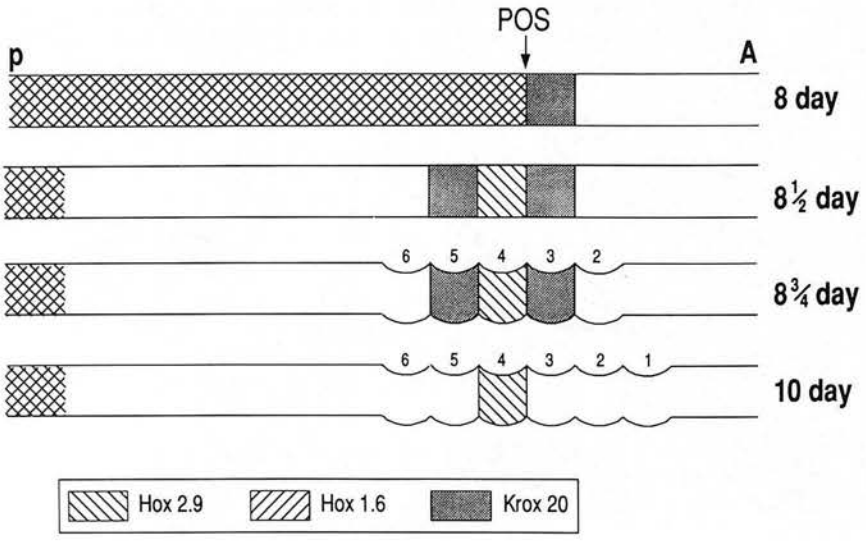


Fig. 9. A diagrammatic representation of the expression patterns of *Hox 2.9*, *Hox 1.6* and *Krox 20* in the CNS during segmentation of the hindbrain. The first appearance of rhombomeres is represented at $8\frac{3}{4}$ days, at which time at least 5 rhombomeres are visible. The diagram does not represent the relative size of the CNS at the various stages.

have the capacity to bind DNA directly, it may maintain the capacity to interact with other regulators and in this way may be involved in a complex regulatory network. Alternatively the splicing mechanism may be involved in removing functional *Hox 1.6* protein without shutting down the transcription of the gene. This paper describes extensive similarities in the expression patterns of *Hox 2.9* and *Hox 1.6* indicating that they may be similarly controlled in the embryo. This offers an explanation for the presence of such an additional control mechanism to down-regulate *Hox 1.6* without affecting *Hox 2.9*. Our results show that the relative amount of homeodomain-producing transcript drops dramatically between 8 and 9 days and that at 9 days it is only detectable by *in situ* hybridisation in the gut epithelium. This would imply that homeobox-containing protein from both genes is required to pattern the 8 day embryo but there is a greater requirement for *Hox 2.9* as a transcription factor at later stages.

We thank Duncan Davidson for much advice, assistance and stimulating discussion and also for critically reading this manuscript. We thank Gillian M. Morriss-Kay for the sections shown in Fig. 7, the photographic department and in particular Sandie Bruce for his time and patience and Liz Graham for technical assistance with *in situ* hybridisation. The *Krox 20* probe was gratefully received from D. Wilkinson.

References

Acampora, D., D'Esposito, M., Faiella, A., Pannese, M., Migliaccio, E., Morelli, F., Stornaiuolo, A., Nigro, V., Simeone, A. and Boncinelli, E. (1989). The human *HOX* gene family. *Nucl. Acids Research* **17**, 10385–10402.

Akam, M. (1987). The molecular basis of metameric pattern in the *Drosophila* embryo. *Development* **101**, 1–22.

Awgulewitsch, A., Utset, M. F., Hart, C. P., McGinnis, W. and Ruddle, F. H. (1986). Spatial restriction in expression of a mouse homeobox locus within the central nervous system. *Nature* **320**, 328–335.

Baron, A., Featherstone, M. S., Hill, R. E., Hall, A., Galliot, B. and Duboule, D. (1987). *Hox 1.6*: a mouse homeobox-containing gene member of the *Hox 1* complex. *EMBO J.* **6**, 2977–2986.

Bogarad, L. D., Utset, M. F., Awgulwitsch, A., Hart, C. P. and Ruddle, F. H. (1989). The developmental expression pattern of a new murine homeobox gene: *Hox 2.5*. *Devl Biol.* **133**, 537–549.

Breier, G., Dressler, G. R. and Gruss, P. (1988). Primary structure and developmental expression of *Hox 3.1*, a member of the murine *Hox 3* homeobox cluster. *EMBO J.* **7**, 1329–1336.

Bucan, M., Yang-Feng, T., Colberg-Poley, A. M., Wolgemuth, D. J., Genet, J., Franke, U. and Lehrach, H. (1986). Genetic and cytogenetic localisation of the homeobox-containing genes on mouse chromosome 6 and human chromosome 7. *EMBO J.* **5**, 2899–2905.

Chavrier, P., Zerial, M., Lemaire, P., Almendral, J., Bravo, R. and Charnay, P. (1988). A gene encoding a protein with zinc fingers is activated during G0/G1 transition in cultured cells. *EMBO J.* **7**, 29–35.

Couly, G. and LeDouarin, N. M. (1990). Head morphogenesis in embryonic avian chimeras: evidence for a segmental pattern in the ectoderm corresponding to the neuromeres. *Development* **108**, 543–558.

Davidson, D. R., Graham, E., Sime, C. and Hill, R. E. (1988). A gene with sequence similarity to *Drosophila engrailed* is expressed during the development of the neural tube and vertebrae in the mouse. *Development* **104**, 305–316.

Diederich, R. J., Merrill, V. K. L., Pultz, M. A. and Kaufman, T. C. (1989). Isolation, structure, and expression of *labial*, a homeotic gene of the Antennapedia complex involved in *Drosophila* head development. *Genes & Dev.* **3**, 399–414.

Dolle, P., Izpisua-Belmonte, J-C., Falkenstein, H., Renucci, A. and Duboule, D. (1989). Coordinate expression of the murine *Hox 5* complex homeobox-containing genes during limb pattern formation. *Nature* **342**, 767–772.

Duboule, D. and Dolle, P. (1989). The structural and functional organisation of the murine *Hox* gene family resembles that of *Drosophila* homeotic genes. *EMBO J.* **5**, 1497–1505.

Duprey, P., Chowdhury, K., Dressler, G. R., Balling, R., Simon, D., Guenet, J-L. and Gruss, P. (1988). A mouse gene homologous to the *Drosophila* gene *caudal* is expressed in epithelial cells from the embryonic intestine. *Genes & Dev* **2**, 1647–1654.

Edgar, B. A., Weir, M. P., Schuibiger, G. and Kornberg, T. (1986). Repression and turnover pattern *fusi tarazu* RNA in the early *Drosophila* embryo. *Cell* **47**, 747–754.

Featherstone, M. S., Baron, A., Gaunt, S. J., Mattei, M-G. and Duboule, D. (1988). *Hox 5.1* defines a homeobox-containing gene locus on mouse chromosome 2. *Proc. natn. Acad. Sci. U.S.A.* **85**, 4760–4764.

Fraser, S., Keynes, R. and Lumsden, A. (1990). Segmentation in the chick embryo hindbrain is defined by cell lineage restrictions. *Nature* **344**, 431–434.

Gaunt, S. J. (1987). Homeobox gene *Hox 1.5* expression in mouse embryos: earliest detection by *in situ* hybridisation is during gastrulation. *Development* **101**, 51–60.

Gaunt, S. J., Krumlauf, R. and Duboule, D. (1989). Mouse homeo-genes within a subfamily, *Hox 1.4, 2.6* and *5.1*, display similar anteroposterior domains of expression in the embryo, but show stage- and tissue-dependent differences in their regulation. *Development* **107**, 131–141.

Gaunt, S. J., Sharpe, P. T. and Duboule, D. (1988). Spatially restricted domains of homeo-gene transcripts in mouse embryos: relation to a segmented body plan. *Development* **104** supplement, 169–179.

Gehring, W. J. (1987). Homeo boxes in the study of development. *Science* **236**, 1245–1252.

Graham, A., Papalopulu, N. and Krumlauf, R. (1989). The murine and *Drosophila* homeobox gene complexes have common features of organisation and expression. *Cell* **57**, 367–378.

Harding, K., Wedeen, C., McGinnis, W. and Levine, M. (1985). Spatially regulated expression of homeotic genes in *Drosophila*. *Science* **229**, 1236–1242.

Hart, C. P., Awgulewitsch, A., Fainsod, A. and Ruddle, F. H. (1985). Homeobox gene complex on mouse chromosome 11: molecular cloning, expression in embryogenesis, and homology to a human homeobox locus. *Cell* **43**, 9–18.

Hart, C. P., Fainsod, A. and Ruddle, F. H. (1987). Sequence analysis of the murine *Hox 2.2, 2.3* and *2.4* homeoboxes: evolutionary and structural comparisons. *Genomics* **1**, 182–185.

Hill, R. E., Jones, P. F., Rees, A. R., Sime, C. M., Justice, M. J., Copeland, N. G., Jenkins, N. A., Graham, E. and Davidson, D. R. (1989). A new family of mouse homeobox-containing genes: molecular structure, chromosomal location, and developmental expression of *Hox 7.1*. *Genes & Dev.* **3**, 26–37.

Holland, P. W. H. and Hogan, B. L. M. (1988). Expression of homeobox genes during mouse development: a review. *Genes & Dev.* **2**, 773–782.

Joyner, A. L. and Martin, G. M. (1987). *En-1* and *En-2*, two mouse genes with sequence homology to the *Drosophila engrailed* gene: expression during embryogenesis. *Genes & Development* **1**, 29–38.

Kaufman, T. C., Lewis, R. and Wakimoto, B. (1980). Cytogenetic analysis of chromosome 3 in *Drosophila melanogaster*: the homeotic gene complex in polythene chromosome interval 84A-B. *Genetics* **94**, 115–133.

Kessel, M. and Gruss, P. (1990). Murine developmental control genes. *Science* **249**, 374–379.

LaRosa, G. J. and Gudas, L. J. (1988). Early retinoic acid

induced F9 teratocarcinoma stem cell gene ERA-1: alternate splicing creates transcripts for a homeobox-containing protein and one lacking the homeobox. *Molec. cell. Biol.* **8**, 3906–3917.

LEWIS, E. B. (1978). A gene complex controlling segmentation in *Drosophila*. *Nature* **276**, 565–570.

LUMSDEN, A. AND KEYNES, R. (1989). Segmental patterns of neuronal development in the chick hindbrain. *Nature* **337**, 424–428.

MERRILL, V. K., DIEDERICH, R. J., TURNER, F. R. AND KAUFMAN, T. C. (1989). A genetic and developmental analysis of mutations in *labial*, a gene necessary for proper head formation in *Drosophila melanogaster*. *Devl Biol.* **135**, 376–391.

MLODZIK, M., FJOSE, A. AND GEHRING, W. H. (1988). Molecular structure and spatial expression of a homeobox gene from the *labial* region of the Antennapedia complex. *EMBO J.* **7**, 2569–2578.

MURPHY, P., DAVIDSON, D. R. AND HILL, R. E. (1989). Segment specific expression of a homeobox-containing gene in the mouse hindbrain. *Nature* **341**, 156–159.

NODEN, D. M. (1975). An analysis of the migratory behavior of avian cephalic neural crest cells. *Devl Biol.* **42**, 106–130.

PROUDFOOT, N. J. AND BROWNLEE, G. G. (1976). Non-coding region sequences in eukaryotic messenger RNA. *Nature* **262**, 211–214.

RUBOCK, M. J., LARIN, Z., COOK, M., PAPALOPULU, N., KRUMLAUF, R. AND LEHRACH, H. (1990). A yeast artificial chromosome containing the mouse homeobox cluster *Hox 2*. *Proc. natn. Acad. Sci. U.S.A.* **87**, 4751–4755.

RUIZ I ALTABA, A. AND MELTON, D. A. (1990). Axial patterning and the establishment of polarity in the frog embryo. *TIG* **6**, 57–64.

SANCHEZ-HERRERO, E., VERNOS, I., MARCO, R. AND MORATA, G. (1985). Genetic organisation of *Drosophila* bithorax complex. *Nature* **313**, 108–113.

SANGER, F., NICKLEN, S. AND COULSON, A. R. (1977). DNA sequencing with chain-terminating inhibitors. *Proc. natn. Acad. Sci. U.S.A.* **78**, 5463–5469.

SCOTT, M. P. AND CARROLL, S. B. (1987). The segmentation and homeotic gene network in early *Drosophila* development. *Cell* **51**, 689–698.

SUNDIN, O. H., BUSSE, H. G., ROGERS, M. A., GUDAS, L. J. AND EICHELE, G. (1990). Region-specific expression in early chick and mouse embryos of *Ghox-lab* and *Hox 1.6*, vertebrate homeobox-containing genes related to *Drosophila labial*. *Development* **108**, 47–58.

TAN, S. S. AND MORRISS-KAY, G. M. (1986). Analysis of cranial neural crest cell migration and early fates in postimplantation rat chimaeras. *J. Embryol. exp. Morph.* **98**, 21–58.

UTSET, M. F., AWGULEWITSCH, A., RUDDLE, F. H. AND MCGINNIS, W. (1987). Region specific expression of two mouse homeobox genes. *Science* **235**, 1379–1382.

WILKINSON, D. G., BHATT, S., CHAVRIER, P., BRAVO, R. AND CHARNAY, P. (1989*a*). Segment specific expression of a zinc finger gene in the developing nervous system of the mouse. *Nature* **337**, 461–464.

WILKINSON, D. G., BHATT, S., COOK, M., BONCINELLI, E. AND KRUMLAUF, R. (1989*b*). Segmental expression of *Hox 2* homeobox genes in the developing mouse hindbrain. *Nature* **341**, 405–409.

(*Accepted 5 October 1990*)